Vol. 46 OCTOBER, 1955 No. 6

The Journal of Educational Psychology

UNIVERSITY OF MICHIGAN
JAN
PERIODICAL READING ROOM

CONTENTS

The Relationship of Immediate Retention of Technical Training Material to Career Preferences and Aptitudes 321
G. R. KLARE, L. M. GUSTAFSON, J. E. MABRY AND E. H. SHUFORD

Personality and Teacher Effectiveness: A Problem in Theoretical Research 330
E. G. GUBA AND J. W. GETZELS

The Relative Stability of Classroom Social Status with Unweighted and Weighted Sociometric Choices 345
NORMAN E. GRONLUND

Sex Differences in Adjustment to College 355
JOHN SUMMERSKILL AND C. DOUGLAS DARLING

The Social Structure of the Classroom Group: An Observational Framework 362
GALE E. JENSEN

Mental Maturity versus Perception Abilities in Primary Reading ... 375
SISTER MARY JAMES HARRINGTON AND DONALD D. DURRELL

Book Reviews 381

Published monthly except June to September
$7.00 a year in U. S. and Pan Amer.; Canada, $7.20; other countries, $7.40.
Single issues $1.20

WARWICK & YORK, INC.
BALTIMORE 2, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879.

UNIVERSITY OF MICHIGAN LIBRARIES

THE JOURNAL OF Educational Psychology

Established 1910

BOARD OF EDITORS*

WM. CLARK TROW, *Chairman*
University of Michigan

STEPHEN M. COREY
Teachers College
Columbia University

J. B. STROUD
State University of Iowa

* Appointed by the American Psychological Association.

EDITORIAL CONSULTANTS

PAUL BLOMMERS
CHESTER W. HARRIS
HAROLD E. JONES
H. H. REMMERS
PERCIVAL M. SYMONDS
MILES A. TINKER
ALEXANDER G. WESMAN
PAUL A. WITTY

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. The JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

INFORMATION FOR CONTRIBUTORS

Manuscripts and communications regarding editorial matters should be addressed to the Chairman of the Board of Editors, Wm. Clark Trow, School of Education, University of Michigan, Ann Arbor, Mich.

THE JOURNAL has set regulations regarding content and style of material published, and these should be observed in the preparation of manuscripts to be submitted.

Tables and graphs.—In keeping with the practice of most professional periodicals of like character, THE JOURNAL now requires authors to bear part of the added cost resulting from the use of tables, formulas, and graphs. Prospective contributors may obtain detailed statement as to

(*Continued on Inside Back Cover*)

THE JOURNAL OF

EDUCATIONAL PSYCHOLOGY

Volume 46 October, 1955 Number 6

THE RELATIONSHIP OF IMMEDIATE RETENTION OF TECHNICAL TRAINING MATERIAL TO CAREER PREFERENCES AND APTITUDES[1]

G. R. KLARE, L. M. GUSTAFSON, J. E. MABRY AND E. H. SHUFORD

University of Illinois[2]

This paper presents a study of the relationship of immediate retention test scores (for a technical training passage) to selected 'career preferences' (expressed interests) and to certain measured aptitudes of airmen in indoctrination training. The data were collected as one aspect of several experiments in the learning and retention of technical training materials, chiefly for the purpose of determining whether the variance of retention test scores attributable to expressed interest was likely to be sufficient to require control of the interest variable. Since previous experimentation had shown that aptitude was likely to be a very important contributor to variance in retention test scores, this relationship as well as that of measured aptitude to career preferences was also investigated.

That interest is of potential significance to the process of learning was recognized as early as the Seventeenth Century (*13*). The interests of the learner have even been made the cornerstone of certain later philosophies of education (*13*). Though it was not until World War I that psychologists and educators began to study interest systematically, since that time the relation of interests

[1] This research was supported in part by the United States Air Force under Contract No. AF 33(038)-25726, monitored by the Personnel and Training Research Center. Permission is granted for reproduction, translation, publication, use and disposal in whole and in part by or for the United States Government.

[2] Dr. Klare is now at Ohio University; Dr. Gustafson at Oklahoma A. and M. College; and Dr. Mabry at VA Hospital, Salt Lake City, Utah.

(including career preferences) to measured ability has been the subject of many articles. This literature is too vast for review here, but certain salient points relative to the study reported in this paper will be mentioned.

The relationship of interests and abilities has been studied by means of two rather distinct kinds of interest measures. The first ('expressed interests') is based on subjects' statements of their interests; the second ('inventoried interests') is based on the use of an interest inventory. The latter is a more refined measure both statistically and semantically. These two kinds of interest measures are in themselves not highly related; Anderson (*1*), Berdie (*2*), Brown (*3*), and DiMichael (*5*), among others, have found that the relationship ranges from low to moderate. In those studies in which correlation coefficients were computed, the values tend to be about +.50.

Both kinds of interest measures have been studied in relation to a variety of measures of ability and achievement. Because of this variety, a comparison of studies is difficult to make; both positive and negative relationships have been found for both kinds of interest indices, the correlations being low to moderate in degree (Strong, *12*). Typical recent examples of studies showing little or no relationship are those of Jackson (*6*), Klugman (*8*), and Townsend (*14*). The studies of Kopp and Tussing (*9*) and Laycock and Hutcheon (*10*) are typical of those showing moderate relationships (correlation coefficients up to about +.50). One of the highest correlation coefficients found was reported by Wesley, Corey and Stewart (*15*), who found that a measure of interest in physical science yielded a correlation of +.66 with success in engineering courses.

METHOD

The data reported here were collected at two separate times but in much the same way. The first study served as a pilot experiment; the second as a check on the results of the first, with a larger number of subjects. In both studies the subjects read a technical passage and then took a test covering the passage. Immediately afterward, the subjects were asked to indicate their preferences for a number of Air Force career areas. Mechanical and clerical aptitude scores were also available for these subjects.

Immediate retention tests.—The tests used were of the four-alter-

UNIVERSITY OF MICHIGAN LIBRARIES

native, multiple-choice form and covered a 1206-word printed lesson on the induction and cooling systems of an aircraft engine. The material was rather highly technical in nature and was one of the series of lessons used in training Air Force mechanics at Chanute AFB.

In the first study, the standard version of this material was read by the subjects and two test forms, called PD and PE, were used. In the second study, the standard version was again used along with several additional versions that contained the same content but differed in manner of presentation of the material. A test called form G, made up of items selected from tests PD and PE, was used with all versions in the second study. Test scores on the versions used did not differ significantly.

Aptitude indices (stanine scores).—Air Force personnel records contain aptitude indices on a 'stanine scale' (from 1 to 9) for all airmen. These indices, based on the Airman Classification Test Battery (*4*), cover aptitudes in eight areas: mechanical, clerical, equipment operation, radio operation, technical specialty, service, crafts, and electronics. In both studies reported here only the mechanical and clerical indices were used in analysis, the former because the reading material was mechanical in nature, the latter because of its relationship to reading ability.

Career preference check list.—The career preference check list covered eight general career areas, with a short description of each, based on the eight aptitude areas covered by the Airman Classification Test Battery. These areas were described as common to civilian life, and subjects were told to indicate with a '1' their first preference, '2' their second preference, etc., down to '8' as the least preferred area. They were told to answer as if they had equal chances for training and experience—and equal chances for pay and advancement—in each area. As in the case of the aptitude data, only the mechanical and clerical preference data were used in analysis.

Subjects.—Subjects for both the first and second studies were airmen in basic training at Sampson Air Force Base, an indoctrination center. Both studies were carried out during the early part of the basic training program, just after the regular classification tests had been administered.

For the first study, there were one hundred fifty-nine subjects who took retention test form PD and one hundred fifty who took

test form PE. For the second study, there were four hundred forty subjects (all took the form G test).

RESULTS

Mechanical career preference and retention test score for a mechanical training lesson.—First study: It will be recalled that the first study was a pilot study, with relatively small N's (one hundred fifty and one hundred fifty-nine) in two test groups (each taking a different test form). It was noted in the first study that the mechanical work area was by far the most preferred of all the career areas listed; there was, in consequence, a relatively large number of subjects who listed mechanical work as a high preference and a relatively small number who listed it as a middle or low preference. In order to get about equal N's for high and low preference groups, high preference airmen were defined as all subjects who gave mechanical work a preference of 1, and low preference airmen as all subjects who gave mechanical work a preference of 6, 7 or 8.

Examination of the retention test score data for the high and low preference airmen showed that the former had somewhat higher test scores than the latter (in the case of both the PD and PE test groups). However, *t*-tests showed that the mean differences were not significant.

Second study: A 3 × 8 disproportionate analysis of variance was computed on the immediate retention test (form G) scores of the four hundred forty subjects. Three levels of mechanical work preference (1–2, 3–5, 6–8) and eight levels of mechanical aptitude (1–2, 3, 4, 5, 6, 7, 8, 9) were used as the independent variables (see Table 1). Though inspection of the data showed that higher preference was related to somewhat higher immediate retention test scores just as in the first study, the variance attributable to mechanical work preference in the analysis was not significant at the .05 level. The variances attributable to mechanical aptitude and to interaction were significant beyond the .05 level, but no consistent pattern of interaction effect was apparent.

Clerical career preference and retention score for a mechanical training lesson.—Analysis of the data for clerical career preferences and retention scores showed there was little relationship between the two in either the first or the second study.

Mechanical aptitude and retention score for a mechanical training lesson.—Table 1 is consistent with other work not reported here in

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 1.—ANALYSIS OF VARIANCE OF IMMEDIATE RETENTION SCORES OF 440 SAMPSON AIRMEN IN RELATION TO MECHANICAL WORK PREFERENCES AND MECHANICAL APTITUDE INDEX SCORES

Source of variation	Sum of squares	df	Mean square	F	P
Mechanical work preferences	15.76	2	7.88	.25	>.05
Mechanical stanine scores	21,879.20	7	3,125.60	99.82	<.05
Interaction	919.14	14	65.65	2.10	<.05
Within	13,023.61	416	31.31		
Total	35,837.71	439			

showing that mechanical aptitude index can be expected to account for significant retention test variance. To ascertain the degree of the aptitude-retention score relationship, several product-moment correlations were computed. For the first study, the correlation coefficients were +.74 for the PD (N = 150) and +.75 for the PE (N = 159) groups. For the second study, the four hundred forty subjects' scores were analyzed as four groups of one hundred ten each (all groups took the Form G test); the correlation coefficients were +.68, +.75, +.78, and +.83. The relationship between mechanical aptitude and retention test scores in these studies thus appears quite high and stable.

Clerical aptitude and retention score for a mechanical training lesson.—The relationship for clerical aptitude and retention test score, though lower than that for mechanical aptitude and test score, is also quite high and stable. For the first study, the correlation for the PD group was +.55 and for the PE group +.60. For the second study, the correlations were +.54, +.55, +.62, and +.67. Intercorrelations of mechanical and clerical aptitude indices for the two studies ranged from +.63 to +.83 (median value, +.78).

Mechanical career preference and mechanical aptitude.—First study: In the first (pilot) study, a 2 × 2 disproportionate analysis of variance[3] (high versus low preference × groups PD and PE)

[3] The method of analysis was that suggested by Snedecor (*11*) for the case of disproportionate frequencies in subclasses. In addition to computing Bartlett's test for homogeneity of variance, the precaution suggested by Johnson (*7*) of computing chi-square to determine whether actual cell frequencies differ significantly from those expected was also followed in all disproportionate analyses.

was computed on mechanical aptitude scores. It was found that the variance attributable to preference was significant beyond the .05 level; the variance attributable to replication (test groups) and that attributable to interaction were not significant. Examination of the relevant means showed that subjects with a high preference for mechanical work had lower aptitude scores than subjects with a low preference.

Second study: In the second study, a 3 × 1 analysis of variance was computed on mechanical aptitude scores, using three levels of mechanical work preference (1–2, 3–5, 6–8). It was found that the variance attributable to mechanical work preference was again significant beyond the .05 level, as in the first study. The same trend of means was also found, the subjects with a higher mechanical work preference having a lower mechanical aptitude score than the subjects with a lower preference. Since the distribution appeared curvilinear, an eta coefficient was computed and found to yield a low value.

Clerical career preference and clerical aptitude.—First study: In order to get about equal N's for the high and low clerical preference groups, the high preference group was defined as all subjects with a preference of 1 or 2 and the low preference group as all subjects with a preference of 8 (clerical work was among the least preferred of the eight career areas). A 2 × 2 disproportionate analysis of variance (high versus low preference × groups PD and PE) was computed on clerical aptitude scores. It yielded variance attributable to preference significant beyond the .05 level; the variance attributable to replication (test groups) and that attributable to interaction were not significant. Examination of the means showed that subjects who had a high clerical preference had somewhat higher clerical aptitude scores than subjects who had a low preference. This relationship, it will be noted, is the opposite of that found for mechanical preference and mechanical aptitude.

Second study: A 3 × 1 analysis of variance was computed on clerical aptitude scores, using three levels of clerical work preference (1–3, 4–6, 7–8). The variance attributable to preference was found to be significant beyond the .05 level, which is in agreement with the results of the first study. Also, as in the first study, subjects with a higher clerical preference were found to have higher clerical aptitude scores than subjects with a low preference. An eta coefficient computed on these data yielded a low value.

UNIVERSITY OF MICHIGAN LIBRARIES

DISCUSSION

These studies show some tendency, albeit non-significant, for higher mechanical preference to be related to higher retention test achievement in the mechanical area. The non-significant relationships in these particular studies may have been due to the observed high motivation of all subjects in the test situation (which may well not hold up through an entire course of training), to a test not quite sensitive enough to show clear differences, or to other reasons. Further work is needed to determine whether the relationships found here are of any general significance.

In contrast to the preference-retention test results, the aptitude-retention test results clearly indicate a relationship that is consistently high. While this is not surprising, it does suggest that in such studies where aptitude indices are available the variance attributable to them should be accounted for in statistical analysis. The magnitude of the relationship indicates that this variance should be considered even when experimental groups appear to differ only relatively slightly in aptitude.

Apart from these results, the studies indicate certain interesting relationships between work preferences and aptitude scores. While the relationships are low—both for mechanical work preference and mechanical aptitude and for clerical work preference and clerical aptitude—the two studies are in agreement in disclosing significant relationships. This is of particular interest since the relationship for the former pair of factors is negative, and for the latter is positive. While both negative and positive relationships have been found in the literature, the observation that most are low suggests that the direction of the relationships may in many cases have been due to chance. The consistency in the two studies reported here, however, suggests somewhat greater stability for the present results.

One possible reason for the low relationship between mechanical work preference and mechanical aptitude is the experience in, and attitude toward, mechanical work that younger males get in this 'machine-age' society. That is, mechanical work is looked upon more favorably by most persons than many service occupations, and most persons have at least some mechanical experience. The relatively high preference status of mechanical work compared to other work areas has already been pointed out. This high preference status seems to indicate that interest in mechanical work

arises largely from factors other than aptitude (or ability) and is fairly well spread over the aptitude range. Thus a low correlation between interest and aptitude in this work area would be expected.

Consideration of preference and aptitude for other work areas suggests, further, why the mechanical preference-aptitude relationship may be negative. Those subjects in these studies who received high mechanical aptitude scores also generally received high aptitude scores in other areas (an example from these studies is the median correlation of +.78 between mechanical and clerical aptitude scores). It seems possible that usually only the more able subjects in other work areas have highest career preferences in those areas (the positive relationship between clerical aptitude and clerical work preference is an example). This would tend to remove many of the subjects with higher mechanical (as well as other) aptitudes and leave a relatively larger proportion of the lower aptitude subjects to choose mechanical work. The low mechanical preference-aptitude relationship would become negative to the extent that this occurs, since it would tend to leave the lower-aptitude subjects with the higher preferences for mechanical work.

SUMMARY AND CONCLUSIONS

The two studies reported here were designed to determine the relationships of score on an immediate retention test covering a mechanical training lesson, selected career preferences (expressed interests), and selected airman aptitude indices. Neither mechanical work preferences nor clerical work preferences showed any significant relationship to retention test scores on the lesson. There was a tendency for higher mechanical work preference to be related to higher retention scores, but the relationship was too small to reach statistical significance. Both mechanical and clerical aptitude indices, on the other hand, appear quite highly and consistently related to retention test scores (this is particularly true for mechanical aptitude).

In these studies, a low negative relationship was noted between preferences for mechanical work and mechanical aptitude score. The relationship between clerical preference and clerical aptitude was also low, but positive. It is interesting to note that two separate studies yield similar results, indicating that the relationships, while low, were nevertheless stable.

REFERENCES

1) Rose G. Anderson. "Subjective ranking versus score ranking of interest values." *Personnel Psychol.*, 1948, 1, 349–355.

2) R. F. Berdie. "Scores on the Strong Vocational Interest Blank and the Kuder Preference Record in relation to self-ratings." *J. appl. Psychol.*, 1950, 34, 42–49.

3) M. N. Brown. "Expressed and inventoried interests of veterans." *J. appl. Psychol.*, 1951, 35, 401–402.

4) J. T. Dailey. *Development of the Airman Classification Test Battery.* San Antonio, Tex.: Human Resources Research Center, Lackland Air Force Base, November 1948 (*Research Bulletin* 48-4).

5) S. G. DiMichael. "The professed and measured interests of vocational rehabilitation counselors." *Educ. psychol. Measmt.*, 1949, 9, 59–72.

6) J. A. Jackson. "A note on the crystallization of vocational interests." *J. soc. Psychol.*, 1947, 26, 125–130.

7) P. O. Johnson. *Statistical Methods in Research.* New York: Prentice-Hall, 1949.

8) S. F. Klugman. "The effect of schooling upon the relationship between clerical aptitude and interests." *J. genet. Psychol.*, 1945, 66, 255–258.

9) T. Kopp and L. Tussing. "The vocational choices of high school students as related to scores on vocational interest inventories." *Occupations*, 1947, 25, 334–339.

10) S. R. Laycock and N. B. Hutcheon. "A preliminary investigation into the problem of measuring engineering aptitude." *J. educ. Psychol.*, 1939, 30, 280–289.

11) G. W. Snedecor. *Statistical Methods.* Ames, Iowa: Iowa State Coll. Press, 1949.

12) E. K. Strong, Jr. *Vocational Interests of Men and Women.* Stanford Univer. Press, 1943.

13) D. E. Super and J. W. Dunlap. "Interest in work and play." In D. H. Fryer and E. R. Henry (Eds.), *Handbook of Applied Psychology.* New York: Rinehart, 1950.

14) A. Townsend. "Achievement and interest ratings for independent-school boys." *Educ. res. Bull.*, 1945, 43, 49–54.

15) S. M. Wesley, D. Q. Corey and Barbara M. Stewart. "The intraindividual relationship between interest and ability." *J. appl. Psychol.*, 1950, 34, 193–197.

PERSONALITY AND TEACHER EFFECTIVENESS: A PROBLEM IN THEORETICAL RESEARCH[1]

E. G. Guba and J. W. Getzels

The University of Chicago

The idea that teacher effectiveness can be identified and predicted by the application of the usual kinds of selection techniques developed by psychologists and others has resulted in a great amount of research. The results are inadequately represented even in such extensive bibliographies as those by Barr (*2*), and Domas and Tiedemann (*3*). Despite a large number of investigations, relatively little more is known now than was known in 1900. Many of the studies have failed to produce any significant results. Many others, especially those dealing with personality variables, have resulted in pedestrian findings, as for example, that pupils learn better from teachers who are kind, cheerful, and sympathetic to pupil problems. Still others, for which statistically significant findings may be claimed, tend to have little practical significance because they involve an unwieldy number of variables.

What is perhaps the most important reason for this failure to produce generally meaningful and useful results has been stated rather succinctly in the two recent reports (*8*, *9*) of the Committee on Criteria of Teacher Effectiveness of the American Educational Research Association, H. H. Remmers, Chairman. The committee's contention is simply that much past research has been conducted in a theoretical vacuum, seeking *ad hoc* solutions to immediate practical problems, with little regard for the long range usefulness of the findings. Moreover, many of the hypotheses which have been investigated have been based upon an oversimplification of the situation and have been tested by means of an inadequate methodology, leading to conclusions which do not make psychological or sociological sense. The committee concludes that only by predicating hypotheses on sound theory can one hope for useful, relevant, and widely applicable findings.

[1] This research was supported in whole or in part by the United States Air Force under Contract No. AF 18(600)-415 monitored by the Air Research and Development Command.

UNIVERSITY OF MICHIGAN LIBRARIES

The application of theory to research in teacher effectiveness, or indeed, to any research, has, as the committee report implies, several distinct advantages. The initial tendency of the worker in a new field of knowledge is to observe. 'Get the facts' is a familiar exhortation. And this is not without justification, for while the researcher's ultimate purpose may be to explain the phenomena which he will observe, he must first know what these phenomena are. Unless this particular researcher is exceptionally sophisticated, however, he is likely to overlook a vital element in constructing his design for observation; namely, that it is humanly impossible to make objective observations of all elements in a phenomenal field; and, further, that it is unlikely that those phenomena which are observed will be accorded appropriate weights. Consciously or unconsciously, the observer makes a selection from among the available phenomena, and then ascribes differential weightings of relative importance to the various elements which he does select. But what criteria does he utilize in making this selection and ascription? Since such criteria will inevitably be applied, is the researcher not better off for having made them explicit before the task of observation begins?

But theory has a greater usefulness than merely to provide the setting in terms of which a research design may be formulated. A theory is more than taxonomic, it is dynamic; it attempts not only to describe but also to explain. And precisely because it contains such dynamic elements, theory leads directly to further hypotheses which may be tested in terms of their congruence with observational data. An adequate theory is thus self-extensive; i.e., as portions of the theory are tested and found to be valid or invalid, new hypotheses suggest themselves which require further investigation, and so on. The findings resulting from the tests of such hypotheses then serve not only the descriptive function but also, and more importantly, the explanatory function.

Finally, theory provides a basis for interpreting those observational data which are collected. Without the benefit of such a framework, two contradictory observations cannot be resolved; there is no *a priori* basis for supposing that one observation is more valid or relevant than the other. Without such a framework, a single observation which is 'significant' must be accepted at face value. Theory acts here in two ways: first, by flashing a red 'caution' signal when data are uncovered which do not make sense

in terms of the theoretical formulation, and, then, by serving as arbiter in the attempt to resolve apparent inconsistencies. This is not to say that theory will always be taken as correct or that observed data which contradicts the theory as wrong. It simply means that when such inconsistencies occur, they will be apparent and further investigation will be indicated.

An opportunity to demonstrate rather sharply the differences between a theoretical approach and the more traditional procedures arose in connection with a broad study of teacher effectiveness carried out by the present authors in a military setting.[2] Although the primary purpose of the study was not to discover relationships between personality and teaching effectiveness, data which became available made it possible to investigate these relationships utilizing a rather large number of instruments. Of these instruments, six are pertinent to the present paper: Three tests of a type which have frequently been used in normative, non-theoretically oriented research, and three others for which it is possible to suggest theoretical links with effectiveness. The purpose of this paper is to describe the application of these six devices in a study of the effectiveness of Air Force officer-instructors, and to indicate the major points of difference in the findings resulting from the two approaches.

PROCEDURES

The setting of the study was the Air Command and Staff School, Air University, Maxwell AFB. This school consists of eleven teaching units, nine of which maintain courses while the remaining two simply serve as lecturer pools. The subject matter is both civilian and military in nature (e.g., law and logistics). Instruction is carried on by two distinct methods, lecturing and discussion, with most instructors doing both kinds of teaching. Two quite different kinds of teaching situations, which might be expected to make differential demands upon the personalities of the instructors involved, were thus available for study.

The sample.—The general sample consisted of two hundred and thirty-four officer-instructors. A varying proportion of this group (N = 156 to 215) was tested with the six instruments in question and, in addition, assisted in the development of effectiveness cri-

[2] A preliminary report of the larger study may be found in (*5*).

teria. The sample approximated very well the total population of instructors both in terms of rank (which ranged from Chief Warrant Officer to Lieutenant Colonel) and assignment to the eleven divisions. A group of sixty-one superiors, each of whom had at least nominal charge of a section of instructors, was also utilized to assist in the development of effectiveness criteria.

The effectiveness criteria.—The effectiveness criteria consisted of various composites of ratings made of instructors by their fellow-instructors (colleagues) and by their superiors. This procedure had validity in the Air University situation for several reasons: (1) a rating system is familiar to military officers and is part of their daily routine, (2) lectures which an AC&SS instructor give must be rehearsed before an audience of fellow-instructors and/or superiors, and (3) mutual visitation of classes is encouraged and, indeed, almost required of AC&SS instructors. The raters thus had opportunity to observe the instructors whom they were rating and implicitly made such ratings almost every day.

Each instructor and superior was required to list in rank order of effectiveness as lecturers and/or discussion leaders (depending upon the kind of teaching done) all of the instructors who served in his division or section. Standard (T) scores were then computed for each instructor to reduce the data from separate courses to comparable terms, and from these scores three composites were derived as follows:

1) Effective and ineffective lecturers. Persons who achieved T scores in excess of 60 (effective) or less than 40 (ineffective) when rated as lecturers by either colleagues or superiors were retained, provided, that persons rated in excess of 60 by one group but less than 40 by the other were omitted.

2) Effective and ineffective discussion leaders. These persons were selected as above.

3) Effective and ineffective instructors in general. The two combined colleague and superior criteria, one for lecturers and one for discussion leaders, described in (1) and (2) above, were further combined by a similar standard scoring process to produce one general criterion.

Statistical tests made upon the resulting criteria indicate that there is a high degree of agreement between the two rating groups (superior and colleagues) as well as among the raters within a

specific rating group in the naming of effective and ineffective instructors. Other tests indicate that the ratings are non-random and made according to a consistent criterion, whatever this may be.

The normatively oriented instruments.—Three well-known standard personality instruments of a type which have often been used in the traditional, normative research were included in the present study: The *Guilford-Martin Inventory of Factors GAMIN*, the *Guilford-Martin Inventory of Factors STDCR*, and the *Allport-Vernon-Lindzey Study of Values*. These instruments are not in any sense intrinsically normative and non-theoretical, of course, but they are often used in this fashion and are so used here. Whatever the value of these instruments may be at the theoretical level for other purposes, there is no reason to suppose that such a heterogeneous collection of variables could meaningfully be synthesized into a theory of teacher effectiveness.

Both Guilford-Martin instruments contain a relatively large number of items to be answered in terms of a Yes-?-No key. Ten scores are available:

1) G—general pressure for overt activity.
2) A—ascendency in social situations as opposed to submissiveness; leadership qualities.
3) M—masculinity of attitudes and interests as opposed to femininity.
4) I—lack of inferiority feelings; self-confidence.
5) N—lack of nervous tenseness and irritability.
6) S—social introversion-extraversion. Shyness, seclusiveness, tendency to withdraw in social situations vs. sociability.
7) T—thinking introversion-extraversion. An inclination to meditative or reflective thinking vs. extraversion orientation of thinking.
8) D—depression. Habitually gloomy and pessimistic mood vs. cheerfulness and optimism.
9) C—cycloid disposition. Strong, emotional fluctuations, tendency toward instability, vs. uniformity of moods, evenness of disposition.
10 R—rhathymia. A happy-go-lucky, carefree disposition vs. an inhibited, over-controlled disposition.

The *Study of Values* is based upon the six personality types originally proposed by Eduard Spranger, each of which assumes one of the following values as the basic motivational force:

1) Theoretical. The theoretical person is predominantly empirical, rational, and critical. He sees the systematization of knowledge as his highest goal.
2) Economic. The economic man is interested primarily in utility.

3) Aesthetic. The aesthetic man is interested in beauty, particularly as developed in form and harmony.

4) Social. The social man has for his highest value the love of people, in the altruistic or philanthropic sense. He is interested in human relationships.

5) Political. The political man sees his highest value in power. Competition, struggle, influence, and renown are characteristic of his motives.

6) Religious. This variable is not religion in the usual sense, but a kind of mysticism in which the individual seeks to relate himself to the universe as a whole. The essential value is unity.

All three standard instruments were scored in accordance with published directions (*1*, *6*, *7*).

The theoretically oriented instruments.—Unfortunately, there is as yet no comprehensive theory of teacher effectiveness which would permit a straight-forward development of hypotheses amenable to test. It is possible, however, to select instruments which, to borrow a term from the Remmers committee, may be 'psychologized', i.e., for which a reasonable connection with effectiveness may be postulated on the basis of relevant psychological theory. It is in this sense that the instruments to be described in this section are theoretically oriented.

Whatever the teacher may teach, it is obvious that the teaching is carried on in the context of an interpersonal setting. It is this factor which, more than any other, accounts for the crucial importance of teacher personality in mediating the teaching-learning process. The teacher cannot force the pupil to learn; what he can do is to produce a situation which the pupil will find conducive to learning. To relieve the teaching process of its affective elements is to reduce it to a sterile, highly intellectualized procedure which the pupil is unlikely to find encouraging.

The establishment of a genuinely good working relationship between teacher and pupil thus requires an interaction not only at the intellectual level but also at the level of personality. Consider then the teacher who is unable to enter into such a relationship. Any one of several modes of behavior may result. The teacher may respond to the everyday frustrations of the classroom with some kind of aggressive reaction. On the other hand, the reaction may be suppressed, leaving the teacher to act not in terms of his own motivations and values but in terms of some imagined standard of how a teacher 'should' act. The second reaction may be worse than the first, for while a hostile teacher may be unpleasant, at

least his true feelings are evident; the latter teacher lacks integrity and hides his true feelings behind a misleading facade.

How can these two reactions, which may be labelled frustration-aggression and reaction-suppression, be measured? Rosenzweig has made an analysis of frustration and has developed an instrument, the *Rosenzweig P-F Study*, which is designed to measure six separate variables found to be associated with frustration behavior (*10*, *11*). When the test is scored[3] in terms of the standardized procedures outlined in the Rosenzweig manual (*12*), six measures associated with these six frustration variables are obtained as follows:

Direction of Aggression.

E—extrapunitiveness, a measure of the extent to which aggression is turned onto the environment.

I—intropunitiveness, a measure of the extent to which aggression is turned by the subject upon himself.

M—impunitiveness, a measure of the extent to which aggression is evaded in an attempt to smooth over the incident.

Type of Reaction.

OD—obstacle dominance, responses in which the barrier occasioning the frustration stands out.

ED—ego defensiveness, responses in which the ego of the subject predominates.

NP—need persistence, responses in which the solution of the problem is emphasized.

The second reaction, reaction-suppression, is not so easily measured. Many of the suppressed attitudes, values, and basic dispositions are, of course, not apparent even to the subject himself. Moreover, the subject is not likely to verbalize publicly even those suppressed attitudes which are conscious. Hence, measurement devices must be used which tap deeper levels than those readily available through conscious recall, and whose purpose is not immediately apparent to the subject.

A suitable technique for the purpose at hand is the method of paired direct and projective questionnaires, devised by one of the authors (*4*). Depending upon the instructions and the item form,

[3] The authors are indebted to Francis J. Van Bortel of the Committee on Human Development, University of Chicago, who scored the Rosenzweig tests.

UNIVERSITY OF MICHIGAN LIBRARIES

it is possible through the use of two such paired instruments to elicit information which bears upon both the overt and the covert attitudes of the subject toward the same object, and hence makes possible the measurement of the disparity between manifest and latent feelings.

Considerable evidence exists (*4, 13*) which shows that subjects respond to two such forms in quite different ways. Typically, the response to the projective instrument is drawn from the deeper layers of personality, and may, especially in the case of a negative response, represent an attitude or feeling which is suppressed under normal circumstances. Responses to the direct instrument may in part, at least, reflect the subject's beliefs about how he 'should' act. To the degree to which the responses to the two forms are similar, the subject is relatively spontaneous and adjusted; while to the extent to which they differ, the subject is relatively repressed and maladjusted.

On the basis of such reasoning, a number of specific hypotheses[4] may be drawn concerning probable relationships between the various test scores and teacher effectiveness:

The *Rosenzweig P-F Study*, Direction of Aggression:

1) Extrapunitiveness, which may be interpreted as the tendency to place blame upon the environment when frustrations occur, is linked to ineffectiveness.

2) Intropunitiveness, which may be interpreted as the tendency to accept responsibility for frustrations that occur, is linked to effectiveness.

3) Impunitiveness, which may be interpreted as the tendency to avoid placing blame altogether, is linked to effectiveness.

The *Rosenzweig P-F Study*, Type of Reaction:

4) Obstacle dominance, which may be interpreted as the tendency to focus attention upon the obstacle occasioning the frustration, is linked to ineffectiveness.

5) Ego defensiveness, which may be interpreted as the tendency to focus attention on self, is linked to ineffectiveness.

6) Need persistence, which may be interpreted as the tendency to focus attention on the solution to the problem, is linked to effectiveness.

---

[4] These hypotheses may break down in the case of a sample which deviates extremely from national norms for one or more of the test variables. Comparison of the present sample with national norms indicates that the vast majority of subjects, including those labelled as 'ineffective', are well within the limits of normality; indeed, most of the present subjects exceed the national norms in the direction of preferred behavior.

The *Paired Direct and Projective Questionnaires*.

7) A high negative score on the projective form is linked to ineffectiveness.

8) A high negative score on the direct form is linked to neither effectiveness nor ineffectiveness, since both groups may have the same stereotypic beliefs about desirable behavior for the teacher.

9) A large disparity between negative scores on the projective and direct instruments (P-D), indicating a lack of spontaneity and integrity, is linked to ineffectiveness.

FINDINGS

For each of the sixteen normatively oriented variables and the nine theoretically oriented variables, comparisons were made between effective and ineffective lecturers, discussion leaders, and instructors in general by means of Fisher's *t* test. The findings for the normative variables are given in Table 1, and for the theoretical variables in Table 2. Both tables indicate the group obtaining the higher mean score for each variable, together with the associated level of significance.

The difference between the two tables which is noted immediately is the relatively greater proportion of significant differences which were uncovered with the theoretical instruments. Of the sixteen normative variables of Table 1, only 1 or 6.3 per cent are significant for lecturers, 6 or 37.5 per cent for discussion leaders, and 2 or 12.5 per cent for instructors in general. On the other hand, of the nine theoretical variables of Table 2, 67.0 per cent are significant for all three groups—lecturers, discussion leaders, and instructors in general. If the criterion of good research were simply the efficient selection of instruments, that is, selection in such way that a maximal number of significant variables were uncovered in proportion to the total number of variables tested, research conducted with theoretically selected instruments would seem to have a marked advantage.

A more useful comparison of the data provided by the two classes of instruments may be made in terms of interpretability, however. Normative instruments are generally selected because of the broad character of the variables under test, while theoretical instruments are selected on the basis of hypotheses developed from the application of relevant psycho-sociological theory. The data of Table 2 are thus immediately interpretable; the only question is whether or not the data tend to confirm or disprove the hypotheses. The

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 1—DIRECTION AND SIGNIFICANCE OF THE DIFFERENCE OF MEAN SCORES ACHIEVED ON SIXTEEN NORMATIVELY ORIENTED PERSONALITY VARIABLES BY EFFECTIVE AND INEFFECTIVE LECTURERS, DISCUSSION LEADERS, AND INSTRUCTORS IN GENERAL

Variable	Lecturers		Discussion Leaders		Instructors in General	
	Group with Higher Mean Score	Level of Significance	Group with Higher Mean Score	Level of Significance	Group with Higher Mean Score	Level of Significance
The Guilford-Martin Inventory of Factors GAMIN						
General Activity	Effective	NS	Effective	NS	Effective	NS
Ascendance	Effective	NS	Effective	NS	Effective	NS
Masculinity	Effective	NS	Effective	.10	Effective	NS
Inferiority (freedom from)	Ineffective	NS	Effective	.10	Effective	NS
Nervousness (freedom from)	Effective	NS	Effective	NS	Effective	NS
The Guilford-Martin Inventory of Factors STDCR						
Social Intraversion	Ineffective	NS	Ineffective	NS	Ineffective	NS
Thinking Intraversion	Effective	NS	Ineffective	NS	Ineffective	NS
Depression	Effective	NS	Ineffective	.05	Ineffective	NS
Cycloid Disposition	Effective	NS	Ineffective	.01	Ineffective	.10
Rhathymia	Effective	.10	Ineffective	NS	Effective	NS
The Allport-Vernon-Lindzey Study of Values						
Theoretical	Effective	NS	Effective	NS	Effective	NS
Economic	Ineffective	NS	Ineffective	.05	Ineffective	NS
Political	Ineffective	NS	Effective	NS	Ineffective	NS
Aesthetic	Effective	NS	Effective	.05	Effective	.02
Social	Ineffective	NS	Effective	NS	Effective	NS
Religious	Ineffective	NS	Ineffective	NS	Ineffective	NS

TABLE 2—DIRECTION AND SIGNIFICANCE OF THE DIFFERENCE OF MEAN SCORES ACHIEVED ON EIGHT THEORETICALLY ORIENTED PERSONALITY VARIABLES BY EFFECTIVE AND INEFFECTIVE LECTURERS, DISCUSSION LEADERS, AND INSTRUCTORS IN GENERAL

Variable	Lecturers		Discussion Leaders		Instructors in General	
	Group with Higher Mean Score	Level of Significance	Group with Higher Mean Score	Level of Significance	Group with Higher Mean Score	Level of Significance
The Rosenzweig P-F Study						
Extrapunitiveness	Ineffective	.05	Ineffective	.10	Ineffective	.05
Intropunitiveness	Effective	.01	Effective	.02	Effective	.02
Impunitiveness	Effective	NS	Effective	NS	Effective	NS
Obstacle Dominance	Ineffective	.10	Ineffective	NS	Ineffective	NS
Ego Defensiveness	Ineffective	NS	Ineffective	.02	Ineffective	.02
Need Persistence	Effective	.05	Effective	.01	Effective	.01
The Sentence Completion Tests						
Projective Negative Score	Ineffective	.10	Ineffective	.01	Ineffective	.01
Direct Negative Score	Effective	NS	Ineffective	NS	Ineffective	NS
Difference Score (P – D)	Ineffective	.02	Ineffective	.01	Ineffective	.02

following conclusions may thus be drawn immediately:

Hypothesis 1: Extrapunitiveness is linked to ineffectiveness. Hypothesis confirmed at the .10 level for discussion leaders, at the .05 level for lecturers and instructors in general.

Hypothesis 2: Intropunitiveness is linked to effectiveness. Hypothesis confirmed at the .02 level or better for all three groups.

Hypothesis 3: Impunitiveness is linked to effectiveness. Hypothesis in abeyance. No significant differences found, although the observed differences are all in the postulated direction.

Hypothesis 4: Obstacle dominance is linked to ineffectiveness. The hypothesis is provisionally confirmed for lecturers at the .10 level, non-significant but in the postulated direction for the other two groups.

Hypothesis 5: Ego defensiveness is linked to ineffectiveness. Hypothesis verified at the .02 level for discussion leaders and instructors in general, non-significant but in the postulated direction for lecturers.

Hypothesis 6: Need persistence is linked to effectiveness. Hypothesis verified at the .05 level for lecturers, at the .01 level for the two other groups.

Hypothesis 7: A high negative projective score is linked to ineffectiveness. The hypothesis is verified at the .10 level for lecturers, at the .01 level for the other two groups.

Hypothesis 8: A high negative direct score is related to neither effectiveness nor ineffectiveness. The hypothesis is verified in that (a) there are no significant differences, and (b) the direction of the differences is reversed for lecturers and discussion leaders.

Hypothesis 9: The difference between projective and direct negative scores is linked to ineffectiveness. Hypothesis verified at the .01 level for discussion leaders, at the .02 level for the other two groups.

Thus, it is clear that there is supportive evidence for each of the *a priori* hypotheses, although this evidence is not always significant in the statistical sense. In this connection, an example of the usefulness of theory in interpreting empirical data may be found in the case of the variable 'impunitiveness'. Hypothesis had indicated that this should be linked with effectiveness, and indeed it was, although not at statistically significant levels. If this variable had been one of several in a normative battery, an appropriate conclusion would have been that there is no relation between impunitiveness and effectiveness. The fact that it was included on a theoretical basis, however, gives one pause before dismissing it as actually non-significant. Are not the trends, weak as they may be, in the postulated direction? Does not this fact lend credence to

the original formulation and suggest that possibly with another better-selected sample, or with some more refined technique, a significant relationship might not be found?

On the other hand, consider the findings reported in Table 1. Interpretation cannot be made simply by comparing findings to hypotheses, for such hypotheses do not exist. The data must then be accepted at face value, and explained purely on the basis of *a posteriori* rationalization. Consider for example the variable 'masculinity'. Why should effective discussion leaders have a high masculinity score, while for lecturers and instructors in general this variable seems to make no difference? To be sure, it may be recalled that the group studied here consists of Air Force officers among whom masculinity may presumably be prized, and that such masculinity may more easily be asserted in the face-to-face discussion situation than in a passive lecturing situation. But this is pure presumption, for which the data provide no test. This may be contrasted with the procedure of making the presumptions beforehand, and then arranging the design so as to make possible a test of their validity.

One final note may be made in connection with Tables 1 and 2. In so far as the data of Table 1 are significant at all, the significant differences tend to occur in connection with discussion leaders rather than lecturers or instructors in general. Since discussion leading obviously depends upon personal interaction to a much greater extent than does lecturing, this finding is not especially surprising. For Table 2, however, variables significant for discussion leaders are also (with one exception) significant for lecturers. If it may be asserted that more sensitive techniques are required to detect the influence of personality in the lecturing than in the discussion situation, it would appear that theoretical instruments directed specifically at particular variables are more successful in detecting such influences than are the more normative instruments.

CONCLUSIONS

On the basis of the analysis made at the beginning of this paper, and as this is supported by the admittedly limited and selected data of Table 1 and Table 2, the following relative advantages of theoretical over fact-finding or normative research in exploring relationships of personality and teaching effectiveness, may be

noted:

1) The development of a theoretical framework sharpens the research objectives and makes specific the criteria of observation which are to be applied.

2) Variables which cannot meaningfully be organized into a psycho-sociological frame of reference are eliminated from the research design before they become wasteful of the researcher's resources.

3) The probability is increased that the research will be more efficient in the sense that a greater proportion of the variables included in the design will prove to be significant.

4) The usually complex task of interpretation is greatly simplified, sometimes involving little more than comparing findings with hypotheses.

5) Variables which are found to be non-significant, if they are in the theoretically postulated direction, are not likely to be eliminated without additional tests. Failure to detect important differences because of failure in methodological technique is thereby minimized.

6) The development of consistent hypotheses provides an opportunity not only for their test but also for the test of the underlying theory. The value of each unit of research is thereby increased.

7) The testing of specific hypotheses rather than a large number of unrelated variables increases the probability that differences which require sensitive handling will be uncovered.

8) Finally, the development of a systematic theory which can be tested at a large number of points will make possible the comparison of data derived from individual studies and will thereby expedite the emergence of psychologically and sociologically sound formulations.

REFERENCES

1) G. W. Allport, P. E. Vernon, and G. Lindzey. *Manual of Directions for a Study of Values*, Revised edition. Boston: Houghton-Mifflin Co., 1951.

2) A. S. Barr. "The measurement and prediction of teaching efficiency: a summary of investigations." *J. Experi. Educ.*, 1948, 16, 202–283.

3) S. J. Domas and D. V. Tiedemann. "Teacher competence—an annotated bibliography." *J. Experi. Educ.*, 1950, 19, 101–218.

4) J. W. Getzels. "The relationship between covert and overt levels of

attitude and personality organization as revealed by the method of paired projective and direct questionnaires." *Amer. Psychol.*, 1952, 7, 318 (abs.)

5) J. W. Getzels and E. G. Guba. "Rôle, rôle conflict, and effectiveness: an empirical study." *Amer. Sociol. Rev.*, 1954, 19, 164–175.

6) J. P. Guilford. *An Inventory of Factors STDCR: Manual of Directions and Norms*, Revised edition. Beverly Hills, Calif.: Sheridan Supply Co., n.d.

7) J. P. Guilford and H. G. Martin. *An Inventory of Factors GAMIN: Manual of Directions and Norms*, First revision. Beverly Hills, Calif.: Sheridan Supply Co., n.d.

8) H. H. Remmers (Chm.) *et al.* "Report of the Committee on the Criteria of Teacher Effectiveness." *Rev. Educ. Res.*, 1952, 22, 238–263.

9) ———. "Second report of the Committee on the Criteria of Teacher Effectiveness." *J. Educ. Res.*, 1953, 46, 641–658.

10) S. Rosenzweig. "The picture-association method and its application in a study of reactions to frustration." *J. Pers.*, 1945, 14, 3–23.

11) S. Rosenzweig, H. J. Clarke, M. S. Garfield, and A. Lehndorff. "Scoring samples for the Rosenzweig Picture-Frustration Study." *J. of Psychol.*, 1946, 21, 45–72.

12) S. Rosenzweig, E. E. Fleming, and H. J. Clarke. "Revised scoring manual for the Rosenzweig Picture-Frustration Study." *J. of Psychol.*, 1947, 24, 165–208.

13) J. J. Walsh. *Socialization Studied by a Sentence Completion Technique.* Unpublished Ph.D. dissertation, University of Chicago, 1953.

THE RELATIVE STABILITY OF CLASSROOM SOCIAL STATUS WITH UNWEIGHTED AND WEIGHTED SOCIOMETRIC CHOICES

NORMAN E. GRONLUND

University of Illinois

The use of the sociometric test, to determine the relative social status of pupils in classroom groups, is finding widespread use among teachers and research workers alike. One of the persistent questions concerning its use, in this regard, is whether or not the choices received should be weighted during the scoring. There are two common practices. One, is to give each choice a value of one regardless of level of choice; the other is to assign arbitrary weights to the various choice levels. In the latter case, the first choice is generally given a value equal to the total number of choices used and each succeeding choice declines in value by one. For example, with three choices a first choice is counted three points, a second choice two points, and the last choice one point. This weighting of choices is based on the assumption that the social significance of sociometric choices declines as the level of choice decreases and the decline in significance is generally proportionate to the weights assigned to the various choice levels. The weights selected, however, are arbitrary and without the support of experimental evidence.

In light of the above contradictory practices regarding the scoring of the sociometric test, it seemed desirable to investigate the problem further. Ideally, this would involve a determination of the social significance of the various choice levels and a comparison of the results with the arbitrary weights commonly assigned. Since this approach would be extremely difficult, if not impossible, it was decided to answer the above question in terms of the stability of the sociometric results. Specifically, this study is an attempt to determine the relative stability of unweighted and weighted social status scores using a sociometric test of three and five choices, with one criterion. It was decided that less than three choices would not give a sufficient spread of scores for a study of their stability and more than five choices would be unnecessary, since it has been shown (*6*) that there is no increase in stability beyond that number. The use of one criterion is in harmony with the com-

mon practice of classroom teachers. Incidental questions arose during analysis of the data and are answered later in this paper.

Although the arbitrary weighting of sociometric choices, in the determination of classroom social status, is a fairly common practice, none of the numerous investigations (*1*, *2*, *3*, *9*) of the stability of sociometric results were concerned with weighted scores. However, both Moreno (*5*) and Northway (*7*) implied the possibility of improving the stability of sociometric results, through weighting, by pointing out that the constancy of sociometric choices declines as the choice level decreases. A review of the literature has revealed no attempt to determine the relative stability of classroom social status based on unweighted and weighted sociometric choices.

METHOD OF INVESTIGATION

The design of the study is relatively simple. A sociometric test, requesting each pupil to select *the five classmates with whom he most prefers to work*, was administered to all of the pupils in nine public-school classrooms[1] on two different occasions. The first test was administered in the middle of January and the second test four months later. There was a total of 216 pupils from the fourth, fifth, and sixth grades included in the study. Of these, 118 were boys and 98 were girls. The nine classes ranged in size from 19 to 30 pupils.

ANALYSIS OF DATA AND RESULTS

The social status of the pupils in each class was determined by four different scoring methods for each administration of the sociometric test. These scoring methods are as follows:

5 choices (unweighted)—1 point for each choice
5 choices (weighted) —5 points for 1st choice
4 points for 2nd choice
3 points for 3rd choice
2 points for 4th choice
1 point for 5th choice
3 choices (unweighted)—1 point for each choice
3 choices (weighted) —3 points for 1st choice
2 points for 2nd choice
1 point for 3rd choice

[1] Grateful acknowledgment is extended to Miss Margaret Henderson, Director of Elementary Education, and the principals and teachers of the Champaign Public Schools for making this study possible.

TABLE I—CORRELATION COEFFICIENTS REPRESENTING THE TEST-RETEST STABILITY OF UNWEIGHTED AND WEIGHTED SOCIAL STATUS SCORES OVER A FOUR-MONTH PERIOD (RHO'S)

Classrooms		5 Choices		3 Choices	
Grade level	N	Unweighted score	Weighted score	Unweighted score	Weighted score
6	28	.85	.83	.82	.81
5	18	.83	.76	.70	.68
5	24	.78	.75	.68	.68
4	30	.74	.75	.72	.71
4	29	.74	.71	.62	.61
4	21	.73	.81	.72	.76
4	20	.73	.80	.69	.83
4	19	.72	.61	.48	.70
4	27	.70	.74	.61	.64
Mean		.76	.75	.67	.71

When the social status of the pupils had been determined for each administration of the test, the relative stability of the results obtained by the four scoring methods, over a four-month period, was determined by use of the Spearman rank-difference method of correlation. The results of this analysis are presented in Table I.

It will be noted in this table that there is no difference in the relative stability of unweighted and weighted choices at either the five-choice or three-choice level. However, inspection of Table I will reveal that the unweighted social status scores with five choices are consistently more stable than the unweighted social status scores with three choices. The difference between means of these two groups was found to be significant at the .01 per cent level. Apparently the social status of pupils is most stable where the results are based on five choices and the above arbitrary weighting of choices does not affect this stability. Since the stability of unweighted and weighted choices is practically identical, it was decided to determine the extent to which the two types of scores were comparable in indicating the social status of pupils. This was done by calculating rank-difference correlation coefficients between the social status of pupils, on the same test, using unweighted and weighted scores. These coefficients ranged from .90 to .97 with a median of .95, where five choices were used, indicating that the relative social status of pupils is essentially the same with both

TABLE II—STABILITY OF VARIOUS SOCIOMETRIC CHOICE LEVELS OVER A FOUR-MONTH PERIOD IN NINE LATER ELEMENTARY CLASSROOMS

Sociometric Choice Level	Per Cent	
	Md	Range
First	72	60 to 86
Second	59	50 to 71
Third	52	38 to 64
Fourth	45	26 to 63
Fifth	38	21 to 47

types of scores. Thus, there appears to be no advantage in spending the additional time required in the calculation of weighted social status scores, where the above weights are assigned.

Stability of Various Sociometric Choice Levels.—In the hope of developing a more refined weighting system that would improve the stability of social status scores, it was decided to analyze the relative stability of the various sociometric choice levels. This involved a comparison of the extent to which the various choices given on the first sociometric test were repeated on the second test, four months later. For example, it was noted on the first test, in each class, which classmembers received the *first choice* given by each pupil. The percentage of these classmembers that were chosen by the same pupils on the second test was then calculated and used as an index of stability for the *first choice*. The same procedure was used to calculate the stability of the other choice levels. Table II summarizes the results of the above analysis. It can readily be seen in this table that there is a declining stability from the first to the fifth choice. The stability of the first choice ranges from 60 per cent, in the class having the lowest stability, to 86 per cent, in the class having the highest stability, with a median of 72 per cent. In contrast, the stability of the fifth choice ranges from 21 per cent to 47 per cent with a median of 38 per cent. From the first to the fifth choice there is a steady decrease in median stability, although there is some overlapping in the range of stability scores at each choice level. Using the relative stability of the various sociometric choice levels as a guide, the following weighting systems were applied to the social status scores and the stability, over a four-month period, was determined for each system of weights.

Weighting System	First Choice	Second Choice	Third Choice	Fourth Choice	Fifth Choice
A	7	6	5	5	4
B	7	6	5	4	3
C	5	4	3	3	2
D	4	3	2	2	1
E	3	3	2	2	1
F	3	2	2	1	1
G	3	2	1	1	1
H	2	1	1	1	1

Of the above weighting systems, the 3–2–1–1–1 weights provided the greatest stability of social status scores. However, the rank-difference correlation coefficients ranged from .70 to .85 with a mean of .77, which is almost identical to the stability coefficients obtained with unweighted choices. Thus, it appears safe to conclude that the stability of sociometric scores, used to indicate the social status of pupils in the classroom, is not improved by the assignment of these arbitrary weights to the various choice levels.

Stability of High and Low Social Status Scores Using Five Unweighted Choices.—Although teachers and research workers are interested in the relative social status of all members of a given group, the extreme social status positions are of special importance. The development and utilization of the leadership potentialities of those pupils with high social status and attempts to improve the social acceptability of those pupils with extremely low social status is of central concern to teachers. Likewise, research workers investigating the correlates of social status direct most of their attention to the positions of leadership and isolation. Consequently, an analysis of the stability of high and low social status scores has practical significance for both groups.

In order to make the high and low social status groups comparable from class to class, Bronfenbrenner's (*3*) fixed frame of reference was applied. This procedure sets nine choices as the lower limit of the high chosen group (stars) and one choice as the upper limit of the low chosen group (neglectees) where five unweighted choices are allotted. On the basis of chance expectancy, these limits are significant at the .02 level.

To determine the stability of the high and low social status groups over the four-month period, the following formula was applied to these categories in each classroom group.

$$\text{Per Cent Agreement} \frac{2 \times \text{Number of Agreements}}{\text{No. in category (Test I)} + \text{No. in category (Test II)}}$$

The four-month stability of the high status group ranged from 44 to 100 per cent with a median of 60 per cent. Since it appeared that many of the pupils in the high chosen group were at the lower limit on one administration of the sociometric test and the loss of one choice would put them out of that category, it was decided to calculate the stability of the high status group with an allowable deviation of one choice below the lower limit. This analysis increased the stability to a range of 67 to 100 per cent with a median stability of 80 per cent.

Applying the above formula to the low status group, the four-month stability ranged from 25 to 100 per cent with a median of 67 per cent. Allowing a one choice deviation, similar to that for the high status group, increased the stability to a range of 57 to 100 per cent with a median stability of 80 per cent. Thus, the stability of the high social status group and the low social status group is comparable, with an average of 80 per cent of the pupils remaining in the same category, or one choice removed from it, over a four-month period. This high degree of stability, compared to the stability coefficients reported in Table I, indicates that the extreme social status scores are more stable than those in the center of the distribution. It also indicates that large fluctuations in social status are relatively rare where five unweighted choices are used.

Stability of Cross-Sex Choices Using Five Unweighted Choices.—Although it has been shown that there is a fairly high degree of stability of social status scores over a four-month period, where five unweighted choices are used, it is suspected that these stability coefficients are reduced, somewhat, by the instability of cross-sex choices. It is common knowledge that there is a cleavage between the sexes during the later elementary period, which in all probability would weaken any sociometric bonds existing between boys and girls. The validity of the above assumption was determined by separately analyzing and comparing the cross-sex choices of boys and girls on the two administrations of the sociometric test. Using the same formula used to determine the stability of the high and low social status groups, it was found that the constancy of boys' choices of girls, in the nine classes, ranged from 10 to 70 per cent with a median stability of 24 per cent. The stability of girls' choices of boys was even lower, with a range of 0 to 67 per cent and

a median stability of 14 per cent. There can be little doubt that the instability of these cross-sex choices reduced the four-month stability coefficients, reported earlier, since approximately one-fifth of the choices allotted in the nine classes were given to members of the opposite sex. Specifically, the median per cent of cross-sex choices on the first administration of the test was 19 for boys and 22 for girls. On the second administration of the test the medians were 17 and 20 per cent, respectively.

DISCUSSION OF RESULTS

An analysis of the above data has revealed that the stability of pupils' social status, over a four-month period, is not improved by the arbitrary weighting of sociometric choices. It is recognized that there may be other reasons for weighting sociometric choices, but until experimental evidence can justify the value of assigning weights, it is suggested that each choice be given a value of one regardless of level of choice.

Since five choices have provided the most stable index of social status, with a single sociometric criterion, it appears advisable to use at least this number of choices where the measurement of social status is being considered. In his original use of the sociometric test Moreno (*5*) allotted two choices for the selection of seating companions. Since that time, choices have ranged from one to an unlimited number with little, or no, justification given for the number selected. Northway (*7*) has used and recommended three choices, but admits there is no experimental evidence to support this number. With the lack of agreement on the number of choices to be used, the feeling has arisen that the number of choices should be adapted to the choosing situation. For example, if committees of four are being formed, three choices is the logical number to allot to each pupil. This practice is based on the assumption that the choosing situation will be more realistic if the number of choices does not exceed the number of available positions in the situation under consideration. However, the reality of the choosing situation can be maintained with five choices by careful phrasing of the directions. In the above illustration, for example, the pupils could be told that five choices are needed since it is usually impossible to give all pupils their first three choices. This is in no way an attempt to mislead the pupils, but rather the solution to a very real problem in sociometric grouping. Besides providing a more stable

index of social status, it will make it easier to organize groups where conflicts arise among the first three choices.

The fairly high degree of stability of social status over a four-month period, indicated in this study, must be interpreted in terms of the grade level and criterion used. Research has shown (*3*, *4*) that the stability of sociometric results, at the elementary level, increases from the first to the sixth grade. Consequently, social status at the early elementary level on the same criterion, would not be expected to attain the four-month stability indicated in this study. On the other hand, stability coefficients calculated solely from fifth- or sixth-grade populations would probably exceed those reported here. Selection of a different criterion would also influence the stability of social status to a certain degree. There is some evidence (*8*) that social status calculated from general criteria, such as the one used in this study, is more stable than that obtained with criteria based on specific skills. It appears that choices on a general criterion may be elicited by less modifiable characteristics related to the social attraction of the total personality, whereas choices on a specific criterion are determined more by situational factors that are constantly changing.

SUMMARY AND FINDINGS

This study was an attempt to determine the relative stability of unweighted and weighted social status scores using a sociometric test of three and five choices. This test, requesting each pupil to select *the five classmates with whom he most prefers to work*, was administered to all of the pupils in nine public-school classrooms on two different occasions, four months apart. An analysis and comparison of the data revealed the following findings.

1) There is *no* difference in the relative stability of social status based on unweighted and weighted sociometric choices, over a four-month period. This held true for both the three- and five-choice levels, using the arbitrary weights most commonly assigned. Rank-difference correlation coefficients between unweighted and weighted scores, on the same administration of the test, ranged from .90 to .97 with a median of .95, indicating that the relative social status of pupils is essentially the same with both types of scores.

2) The stability of the various choice levels on the sociometric test indicated that the first choice was most stable, with a steady

UNIVERSITY OF MICHIGAN LIBRARIES

decrease in stability to the fifth choice. The median stability results were 72 per cent for the first choice; 59 per cent for the second choice; 52 per cent for the third choice; 45 per cent for the fourth choice; and 38 per cent for the fifth choice. Several weighting systems based on the relative stability of the various choice levels produced no increase, above that of unweighted choices, in the stability of social status over a four-month period.

3) The stability of social status based on five unweighted choices was consistently higher than the stability of social status based on three unweighted choices. The mean rank difference correlation coefficients for the two groups were .76 and .67, respectively. This difference between means was significant at the .01 per cent level.

4) The high and low social status scores, using five unweighted choices, were very stable over a four-month period. Using Bronfenbrenner's fixed frame of reference, it was found that of those pupils in the high and low social status categories on the first test, 80 per cent remained in the same category, or one choice removed from it, four months later.

5) The stability of cross-sex choices, over a four-month period, varied from class to class, but in general was relatively low. The stability of boys' choices of girls ranged from 10 to 70 per cent with a median of 24 per cent. The girls' choices of boys were even less stable with a range of 0 to 67 per cent and a median of 14 per cent. Since approximately one-fifth of the choices allotted in the nine classes were cross-sex choices, it was indicated that these choices contributed disproportionately to a reduction in the stability of the social status scores over a four-month period.

A discussion of the results leads to the suggestion that five unweighted choices be allotted to each sociometric criterion where the classroom social status of pupils is being determined.

REFERENCES

1) Merl E. Bonney. "The Relative Stability of Social, Intellectual, and Academic Status in Grades II to IV, and the Inter-Relationships Between These Various Forms of Growth." *Journal of Educational Psychology*, 34 (February, 1943), 88–102.

2) Merl E. Bonney. "The Constancy of Sociometric Scores and Their Relationship to Teacher Judgments of Social Success, and to Personality Self-Ratings." *Sociometry*, 6 (November, 1943), 409–424.

3.) Urie Bronfenbrenner. *The Measurement of Sociometric Status, Struc-*

ture and Development. Sociometry Monographs, No. 6. New York: Beacon House, 1945. Pp. 80.

4) John E. Horrocks, and Mae E. Buker. "A Study of the Friendship Fluctuations of Preadolescents." *Journal of Genetic Psychology*, 78 (June, 1951), 131–144.

5) J. L. Moreno. *Who Shall Survive?* Washington, D. C.: Nervous and Mental Disease Publishing Co., 1934. Pp. 440.

6) W. I. Newstetter, M. J. Feldstein, and T. M. Newcomb. *Group Adjustment: A Study in Experimental Sociology*. Cleveland: Western Reserve University, 1938. Pp. 154.

7) Mary L. Northway. *A Primer of Sociometry*. Toronto, Canada: University of Toronto Press, 1952. Pp. 48.

8) Mary L. Northway. *Appraisal of the Social Development of Children at a Summer Camp*. Toronto, Canada: University of Toronto Studies, Psychology Series, Vol. V, 1940. Pp. 62.

9) S. L. Witryol, and G. G. Thompson. "An Experimental Comparison of the Stability of Social Acceptability Scores Obtained with the Partial-rank-order and Paired-Comparison Scales." *Journal of Educational Psychology*, 44 (January, 1953), 20–30.

UNIVERSITY OF MICHIGAN LIBRARIES

SEX DIFFERENCES IN ADJUSTMENT TO COLLEGE

JOHN SUMMERSKILL and C. DOUGLAS DARLING

Cornell University

A series of studies is being conducted at Cornell University on circumstances associated with good and poor adjustment in college students. These studies are focused on the characteristics of entering freshmen as related to subsequent performance in college. The subjects are freshmen who entered the University in September, 1948. McCormick (*12*) and Thompson (*19*) have reported on academic success and failure among students of arts and science in this class; Bentson (*1*) studied athletic success and failure among men in this class; and a report is in press (*18*) on the over-all incidence of medical disorders among these students. The present study concerns sex differences in performance demonstrated by the students of this class.

The extensive literature on psychological sex differences suggested that sex bears an important relationship to collegiate performance. Johnson and Terman have generalized from considerable evidence to sex differences in social attitudes and emotionality. Scheinfeld's "Women and Men" (*16*) describes, with experimental and anecdotal data, the biological and cultural determination of widespread sex differences. Freeman and Miles (*7*) review studies which demonstrate "significant differences in general interests, attitudes, pursuits, and occupations." A review by Schaffter and Woody (*15*) cites opinion and data on sex influences upon academic success and social status. Kinsey's second volume (*10*), based on a population which is predominately college educated, argues strongly for psychological differences associated with biological sexual expression. Recent large-scale surveys (*6*, *17*) have shown differences in the attitudes of male and female students towards college as an institution; and one comprehensive survey (*8*) found that such attitudinal differences are present and important beyond the undergraduate years.

It is clear that men and women display differences in percepts, thoughts, and feelings, when queried about themselves, other persons, or things. Sex differences in the organization of western

society are clearly reflected in statistics on jobs and wages, outstanding artists, inventors, etc. Less clearly established are the nature and extent of differences in day-to-day behavior of men and women, e.g., success and failure in daily work, illnesses and injuries, use of leisure hours, etc. The writers believe that conclusions about the relative adjustments of men and women will be clarified by continuing research on performance variables of this kind.

METHOD

This study is based on the individual records of the entire class of 1,818 freshmen who entered Cornell University in September, 1948, for the regular four- or five-year undergraduate program. The data for the study included all available records pertaining to the college careers of these students through the succeeding four years, i.e., until June, 1952. These data were obtained from the following sources: office of the University registrar, offices of the various colleges at Cornell, Student Medical Clinic, University Athletic Association, Student Conduct Committee, students' yearbook for 1952. All data were coded and punched on separate IBM cards for each student. Precautions were taken so that confidential data would be accessible only to the investigators.

Cards were sorted into various performance categories as described below. Numbers of men and women in performance groups were tabulated, compared, and evaluated statistically by chi square tests.

RESULTS

There were 402 women among the 1,818 freshmen who matriculated in 1948, ie., 78% of the class was male, 22% female.

Attrition.—Findings for the class as a whole show that 727 out of 1,818 students had left the University without graduating by the spring of 1952. That is, the net attrition rate was 40%. Among those who left, there were clear-cut scholastic reasons ("busted out") in about 25% of the cases; and no apparent scholastic reasons existed (no "probation" or "warning") in about 50% of the cases. Scholastic factors may or may not have been important in the remaining cases who withdrew after having been on academic probation.

There were no differences between male and female students in

rate of attrition. Women accounted for 22% of the class in 1948 and for 21% of the class in 1952. However, only 11% of the 175 outright failures were women, whereas women comprised 38% of the group of 386 students who left without a history of scholastic difficulty. The difference between the two groups represented by the latter percentages is statistically significant (p value from chi square less than .01).

Academic Achievement.—In the class there were 136 students who had been placed on academic probation in two or more different semesters during their four years at the University. This group of inferior scholars was compared with a group of superior scholars who stood in the top tenth of their respective schools according to cumulative grade averages at the end of the junior year. Academic achievement was assessed in this manner to avoid the influence of different grading systems found in different colleges of the University.

The percentage of women among inferior scholars was found to be disproportionately low (5%) and the percentage of women among superior scholars was found to be disproportionately high (30%). These figures represent a difference which is significant at the .01 level of confidence.

Campus Activities.—This class contributed 84 leaders (president or chief officer) to the major campus organizations, fraternities, and sororities. The student yearbook revealed another group of 89 students who did not report participation in a single club, society, committee, or athletic team, during four years at the University.

Women were represented in 20% of the leadership posts, i.e., were represented according to population. However, only 8% of the non-participants were women. The latter represents a significant deviation from the number of women in the class (p value from chi square less than .05).

Health.—The health data showed that 95% of this class visited the Student Medical Clinic during their college careers. On the average, each student experienced six different medical disorders (illness or injury) while at college. There were 126 students with no more than two medical complaints during the full four-year period. There were 112 students with at least 15 complaints during this same period.

The group which had relatively good health records was 17% female. The group with the relatively poor health records was

29% female. This sex difference is statistically significant at the .05 level of confidence. The data indicate that the difference holds for emotional as well as physical illness although numbers in the former category were too small for statistical evaluation of differences.

DISCUSSION

Significant differences in the performance of men and women were observed in each of the areas investigated, ie., academic progress, participation in campus activities, and health status. This indicates that there is a generalized sex difference in adjustment to college, but the direction of the difference depends upon the specific area under consideration. Thus, it is the women who tend more to poor health records, but the men who tend more to poor scholastic records and to non-participation in campus activities. Disrupted performance is found in both men and women; the differences are in the kinds of performance disrupted. This fact is noteworthy because standardized tests have consistently led to the generalization that women are less emotionally stable than men (*9*).

The differences reported here are statistically significant but not large. It is likely that sex accounts for only a minor part of the variation between individuals in the areas studied. In the present-day college environment sex appears to modify rather than critically determine behavior in these areas.

With regard to academic progress, our findings are in agreement with those reported by McNeeley (*13*). He found that fewer women leave college due to academic failure and more leave for "intangible causes." Among these were included "such terms as 'unable to concentrate on work', 'lack of interest', and 'inability to coordinate efforts'." These statements imply a sex difference in motivation for completing college which is also reflected in questionnaire data obtained at Cornell (*17*) and elsewhere (*6*). These surveys showed that more men undergraduates are vocationally oriented in their reasons for attending college. It is likely that fewer women assign "college degree" the priority assigned by men in striving for personal security. Some women undoubtedly withdraw from college simply because other avenues to personal security gain in priority and one cannot assume that withdrawal represents failure or regress in personal adjustment.

Marriage is frequently mentioned as a goal which replaces "college degree" and consequently accounts for the withdrawal of many women from college. The evidence on this is still equivocal. Few women explicitly say that they go to college primarily to further marriage plans (*17*). Nor have studies on attrition detected a substantial number of women explicitly leaving college to be married (*4, 5, 18, 20*). Further, the number of unmarried women among college graduates is known to be disproportionately high (*8*), but we do not know whether this holds for those who leave before graduating. Further research on this topic is needed.

With regard to academic achievement, our findings on failures, probations, and high standing, all favor women students at Cornell. The fact that the number of men admitted to this class was four times the number of women raises the definite possibility that the women freshmen were, on the average a more select group of scholars. In any event, earlier prejudices about women's inability to handle college work are flatly contradicted by these and similar reports (*6, 15*).

The finding that comparatively few women fail to participate in campus activities may also be partly explained by the local situation. All women at Cornell live in dormitories or sororities with comprehensive recreational and social programs. A substantial number of men in this class lived off-campus after their freshman year, and many men had part time employment off-campus. Also, veterans (15%) and married students (3%) were predominately men and their recreational and social activities were, in all likelihood, less closely identified with the University. The central fact, however, is that co-education clearly extends to organized student activities on the present-day campus.

The obtained sex differences in health status cannot be interpreted with finality. A higher morbidity (and lower mortality) rate among women has been reported for non-college populations (*2, 3*). Whether this represents a true difference in the incidence of somatic disease (*16*) or reflects a difference in motivation (*11*) is not known. It is possible that more women visit the Medical Clinic at Cornell because all women live in University buildings with a well-knit counseling system which places emphasis on preventive health measures. However, evidence from a recent study of mononucleosis (*14*), a disease which is relatively prevalent and troublesome in the college-age group, indicates that the incidence

is truly higher among women and that physical signs accompany the more frequent subjective complaints of women students. In any event, there is little doubt that the students who averaged more than two medical complaints each semester throughout their college careers were unduly handicapped in their day-to-day existence. A study focused on the men and women in this category would shed light on psychodynamics correlated with this sex difference.

SUMMARY

Records pertaining to the college careers of 1,818 students in the class entering Cornell University in 1948 were analyzed. Seventy-eight per cent of the class was male, 22 per cent female. When the two groups were compared, the percentages of women among outright scholastic failures, inferior scholars, and non-participants in campus activities, were found to be disproportionately low. The percentages of women among non-academic withdrawals from the University, superior scholars, and students in poor health, were found to be disproportionately high. Findings are interpreted in the light of previous research on sex differences in collegiate performance.

REFERENCES

1) Thomas G. Bentson and John Summerskill. "The Relation of Personal Success in Intercollegiate Athletics to Certain Aspects of Personal Adjustment," *Res. Quarterly*, In press.

2) Antonio Ciocco. "Sex Differences in Morbidity and Mortality," *Quart. Rev. Biol.*, 1940, 15, 59–73, 190–210.

3) S. D. Collins. "Cases and Days of Illness Among Males and Females," *U. S. Publ. Health Serv. Report*, 1940, 55, 47–93.

4) Jean Failing. *A study on attrition among freshmen who entered the N. Y. State College of Home Economics, 1938–1947*. Mimeographed report, Cornell University, 1951.

5) Daniel D. Feder. "Student Population," in *Encyclopedia of educational research*, Walter S. Monroe (Ed.), New York: The Macmillan Co., 1950.

6) Norman Frederiksen and W. B. Schrader. *Adjustment to College*, Princeton: Educ. Testing Service, 1951.

7) Frank S. Freeman and Catherine C. Miles. "Sex Differences," in *Encyclopedia of educational research*, Walter S. Monroe (Ed.), New York: The MacMillan Co., 1950.

8) Ernest Havemann and Patricia West. *They Went to College*, New York: Harcourt, Brace and Co., 1952.

9) Winifred B. Johnson and Lewis M. Terman. "Some Highlights in the

Literature of Psychological Sex Differences Published Since 1920." *J. Psychol.*, 1940, 9, 327–336.

10) A. D. Kinsey, W. B. Pomeroy, C. E. Martin, and P. H. Gebhard. *Sexual behavior in the human female.* Philadelphia: W. B. Saunders. 1953.

11) Frederick H. Lund. "Adolescent Motivation: Sex Differences." *J. Genet. Psychol.*, 1944, 64, 99–103.

12) Jane Ellen McCormick. *Individual factors associated with academic success.* Unpublished master's thesis, Cornell Univ., 1954.

13) John H. McNeeley. *College Student Mortality.* Bull. 11, U. S. Dept. Interior, Office of Educ., U. S. Govt. Printing Office, Washington, 1938.

14) V. R. Mankad. "Infectious Mononeucleosis: An Analysis of 399 Hospital Cases," *Student Med.*, In press.

15) Dorothy Schaffter and Thomas Woody. "Education of Women," in *Enclyclopedia of educational research*, Walter S. Monroe (Ed.), New York: The MacMillan Co., 1950.

16) Amram Scheinfeld. *Women and Men.* New York: Harcourt, Brace and Co., 1944.

17) Social Science Research Center, Cornell Univ. *Report on the Cornell student body*, Mimeographed, 1951.

18) John Summerskill. "Health Services in American Colleges and Universities: Some Questions Answered." *Student Med.*, In press.

19) Martha Margaret Thompson. *Admission information as predictors for graduation.* Unpublished master's thesis, Cornell Univ., 1953.

20) Ruth G. Weintraub and Ruth E. Sally. "Graduation Prospects of an Entering Freshman." *J. ed. Res.*, 1945, 39, 116–126.

THE SOCIAL STRUCTURE OF THE CLASSROOM GROUP: AN OBSERVATIONAL FRAMEWORK

GALE E. JENSEN

Institute of Education, Grinnell College

For roughly forty years we have been studying the instructional designs used by or "on" classroom groups to determine their effectiveness as measured by individual achievement and group productivity. Some of these studies have recognized indirectly that the social structure of the classroom in some way influences group performance and individual learning.[1] Some, through the use of sociometric techniques, have studied certain aspects of the social structure directly and have pointed out its influence upon classroom performance and learning.[2] However, little has been done to provide a framework of concepts that could be utilized to formulate and analyze systematically the classroom problems that are in some way related to the social structure of the classroom group.[3] The purpose of this paper is to outline such a conceptual framework.

This framework, it is hoped, will represent a step toward recognizing that the systematic and conscious utilization of conceptual schemes in educational practice is more powerful for dealing with educational problems than are present "common sense" approaches. The framework is not to be taken as "true" or as the only one inferable from existing research data, but rather accepted as something which, while tentative, is necessary for the orderly examination and description of problems of social structure in the classroom. This, of course, means that the framework is subject to reconstruction whenever more adequate concepts can be formulated.

---

[1] Cf. W. S. Monroe, "Methods of Teaching," *Encyclopedia of Educational Research*. New York: The Macmillan Co., 1950, pp. 745–752.

[2] H. H. Jennings, *Sociometry in Group Relations*. Washington, D. C.; American Council on Education, 1948.

[3] Cf. Ronald Lippitt, "Applying New Knowledge about Human Relations in Educational Practice," *Baltimore Bulletin of Education*, March–June, 1952, pp. 23–30.

THE CONCEPTUAL FRAMEWORK

The form which a description of the social structure of the classroom group takes should be such that it eventually can be integrated with a descriptive framework for examining the dynamics of the classroom. An account, therefore, of the structural aspects of the classroom group should be stated in terms of functional dimensions[4] along which members relate to one another in their attempts (1) to maintain an optimum balance between possible gratifications and deprivations of individual personality needs, and (2) to participate as members concerned with problems both external to the group. Members keep changing their relationships to one another, but always, it is proposed, the changes are along one or more of the dimensions that will be identified. In brief, the dimensions identified indicate all the possible types of functional relationships that individuals can establish between one another *as group members.*[5]

As revealed by the research references cited throughout this article, the dimensions of the social structure of the classroom group represent manifestations of *needs* class members have (a) as individual personalities and (b) as group members. To participate effectively in the tasks or work of the group and to meet their individual needs, members are required to relate to other members in ways that make these accomplishments possible. Stated otherwise, each group member has *need* to establish between himself and the other group members, the kind of relationships that are required for members (1) to work effectively as a group and (2) to maintain a satisfactory balance between the possible gratifications and deprivations of personal needs. These are the basic conditions for a collection of individuals to become and remain a classroom group.

The dimensions along which members seek to relate to one another in a classroom group are seven in number. These are (1)

---

[4] For a more complete statement of the "functional dimensions approach" to social psychology, cf. Jay M. Jackson, "Analysis of Interpersonal Relations in a Formal Organization," unpublished doctoral dissertation, University of Michigan, 1953, pp. 13–17. To be published as a Research Center for Group Dynamics monograph. The theoretical development which follows has been influenced by this work.

[5] Cf. Dorwin Cartwright and Alvin Zander. *Group Dynamics: Theory and Research*. Evanston, Illinois: Row, Peterson and Company, 1953, p. 417.

the problem-solving dimension, (2) the authority-leadership dimension, (3) the power dimension, (4) the friendship dimension, (5) the personal prestige dimension, (6) the sex dimension, and (7) the privilege dimension.

The Problem-Solving Dimension.—There are two kinds of problems to which the classroom must continuously devote its energies.[6] One kind pertains to the efforts which the group will expend to organize relations with its external environment. This constitutes the public (or publicly recognized) task of the group. The other kind of problem refers to the efforts of the group to create the kind of intra-group member relationships (its internal socio-psychological environment) that will satisfy members to such a degree that the group will have sufficient energies for working on problems arising from relationships between itself and its external environment. These two problems, common to all classroom groups, can best be understood through descriptions of the environmental settings to which they are indigenous.

The *external environment* of the classroom group can be depicted as a field of forces that impinge upon the group and require it to expend some of its time and energies coping with the "outside." These forces at any given time may be psychological, economic, political, psychical, or sociological in character.[7] They are the result of conditions existing between two distinct parts of the external environment of the classroom group. These two parts are identifiable as follows. First, the "school" as a social system encompasses or incorporates the classroom group.[8] As a social system, the "school" represents one part of the external environment of the classroom group. Second, the classroom group is basically a training group for enabling people to become competent for participating in the affairs of their physical, biological, and social worlds. Therefore, the other part of the external environment to which the classroom group must relate is that which lies outside the school and which we generally identify as the "community." The tasks, projects, and problems upon which the group works are to a large

---

[6] Herbert Thelen, "Educational Dynamics: Theory and Research," *Journal of Social Issues*, Volume 6, No. 2, 1950, pp. 31–32.

[7] Cf. Clyde B. Moore and William E. Cole. *Sociology in Educational Practice*. New York: Houghton Mifflin, 1952, pp. 33–54.

[8] Gale E. Jensen, "The School as a Social System," *Educational Research Bulletin*, Ohio State University, February, 1954, pp. 38–46.

extent determined by and selected from this part of the external environment. The content of the school curriculum is composed of our systematized knowledges and perceptions of the phenomena and problems that exist in the "community." These systematized knowledges and perceptions (cultural artifacts) about community phenomena and problems are referred to as school-subjects or disciplines.[9] They are utilized by the school as instructional resources or means for achieving the public goals of the school, *viz.*, to teach school participants (students) to live intelligently and satisfyingly in the community.

The *internal environment* of the classroom group can be presented as the inter-member relations operating within the classroom. The group must expend some of its energies in examining the relationships to be organized and rebuilt from time to time between group members. Because these inter-member relations affect the needs systems of individual group members, the problems they create within the classroom group will usually take priority over any others.[10] Little work on problems induced by forces of the external environment will be accomplished when internal difficulties are numerous and strong.

If the classroom group is to contribute effectively to the realization of the public goals of the school, it must deal skillfully with the two ever-present or persisting problems that arise in its internal and external environments. Members of the classroom group must develop perceptions and expectations about what is required in one another's behavior for deriving adequate solutions to the problems. They must make certain that the functions attached to "student" and "teacher" rôles[11] for purposes of dealing effectively with group achievement and group maintenance are activated at the right time. In brief, individuals have, as members of the classroom group, a *need* to relate to one another along this dimension or the class will be unable to establish the group stability and unity required for productive class work and individual development. Failure of class members to relate to one another along this dimen-

[9] Clyde B. Moore and William E. Cole, *Op. Cit.*, pp. 330–331.

[10] N. T. Fouriezos, M. L. Hutt, and H. Guetzkow, "Self-Oriented Needs in Discussion Groups." In D. Cartwright and A. Zander, *Op. Cit.*, pp. 354–360. Max Corey, "Interpersonal Relations and the Work of the School," *Baltimore Bulletin of Education*, March–June, 1952, pp. 16–22.

[11] Cf. Kenneth Benne and Paul Sheats, "Functional Roles of Group Members," *Journal of Social Issues*, Spring, 1948, pp. 42–47.

sion frustrate the strong personal needs of some members to learn more about the area of study for which the class is responsible. Under such conditions, these members are likely to withdraw or engage in non-constructive behavior.

In summary, the ways in which classroom group members relate to one another for working on problems induced by the forces of the external and internal environments of the group, represents the problem-solving dimension of the group. The question which reveals the status of this dimension at any one time is: Who engages in or carries on what problem-solving functions with whom to accomplish the public task of the classroom group?

The Authority-Leadership Dimension.—As a classroom group works on its tasks, it is confronted periodically with having to make decisions that affect its progress. The above discussion referring to the problem-solving dimension of the classroom group indicates that the effectiveness with which these decisions are likely to be made is dependent upon having the problem-solving functions activated at the "right" time. To make certain these problem-solving functions are performed at the right time, the group must appoint and/or recognize some members as having the responsibility for and the power to facilitate the problem-solving conditions requisite to making periodic decisions affecting group progress.[12] The status of this authority-leadership dimension in a class for a given time is revealed by the following question. Who is recognized by whom as having the responsibility and power for facilitating the problem-solving conditions necessary for effective decision-making by the class?

By way of example, if the task on which a class is working should change radically, the class will need to consider changes in on-going and contemplated action. In most instances, one finds that some member of the class is recognized as the responsible person for bringing the class together and calling its attention to certain difficulties with which it may be confronted. This member may carry the designation of "teacher", "chairman", or "leader". He is recognized by class members as the "group convener". He has the power to reconvene the class at the proper time and in a place known to all members.

There are numerous other authority-leadership functions that

[12] Herbert Thelen, "The Experimental Method of Classroom Leadership," *The Elementary School Journal*, October, 1952, pp. 76–85.

must be performed to insure the conditions necessary for and conducive to effective decision-making by a group. These functions must be identified by the class and then assigned to members who will assume the responsibility for facilitating their performance at the crucial points of class operation. Along with identifying and assigning the responsibilities for performing authority-leadership functions, the class must identify and assign, either explicitly or implicitly, the prerogatives to be used in carrying out these functions.[13]

To summarize, the members of the classroom group have a *need* to recognize certain members as having the authority and responsibility for facilitating the problem-solving conditions required for effective decision-making at critical points of class work. In terms of the social structure of the classroom group, this means that class members must establish between themselves a network of authority-leader relationships commensurate with the character of the goals and problems with which the group is confronted.

The Power Dimension.—During the operations of the classroom group it becomes apparent that some members have greater ability (1) to influence the direction the class may take and (2) to induce other members to change their thinking and attitudes. The development of these power relationships in the classroom group is inevitable because members can meet their respective needs only through participation in social situations in which varying needs and resources are matched in ways satisfactory to the participating parties.[14] The character of the relationships along this dimension at any given point in class operation can be discovered by asking: Who possesses what means for gratifying or depriving whose needs?

At times this greater ability to influence others may rest with those who are recognized as having certain authorities. At other times the greater degree of influence is held by those who held no official authority positions. But whether or not this ability to influence rests with those who hold official authority positions is of little importance for the discussion of this dimension. The significant thing is that class members with greater influence are per-

---

[13] Herbert Thelen, "Engineering Research in Curriculum Building," *Journal of Educational Research*, April, 1948, p. 582.

[14] Herbert Thelen, "Basic Concepts in Human Dynamics," *Journal of the National Association of Deans of Women*, March, 1952, p. 104.

ceived at any given time to have greater resources for gratifying the needs of the group and of individual members.[15] It is this ability to gratify or deprive the needs of others that represents the power an individual member possesses in the class. It is this power as manifested in his ability to gratify or deprive needs that enables a member at any given time to influence the direction the group will take or to induce another member to change his perceptions and attitudes.[16]

This power to gratify and deprive needs may be derived from the crucial position a member holds within the group, or it may be derived from his superior ability as an individual member to perform certain acts, control resources, or solve problems. A particular position in the social structure of the class, particular skills, and particular resources, however, are secondary for determining the power a class member holds. Basically, it is the character of the relationships between his social position, resources and skills, and the needs of other members of the class that determines the power a member may have.

To summarize briefly, the way in which class members align themselves with one another for fulfilling individual and group needs through complementary usage of one another's abilities, social positions, and resources represents the power dimension of the classroom group.

The Friendship Dimension.—In the various situations in which class members find themselves, each develops perceptions and feelings about his ability to insure a satisfactory gratification-deprivation balance. Periodic assessments of group situations are made by individual members to determine the possible gratifications and/or deprivations of needs. These assessments are made particularly when the class is trying to project action and/or come to a decision.[17]

---

[15] Ronald Lippitt, Norman Polansky, Fritz Redl, and Sidney Rosen, "The Dynamics of Power." In D. Cartwright and A. Zander, *Op. Cit.*, pp. 480–481.

[16] J. I. Hurwitz, A. Zander, and Barnad Hymovitch, "Some Effects of Power on the Relations among Group Members." In D. Cartwright and A. Zander, *Op. Cit.*, pp. 491–492.

[17] Herbert Thelen, "Educational Dynamics," *Op. Cit.*, pp. 19–21, and H. H. Jennings, *Leadership and Isolation*. New York: Longman's, Green & Co., 1950, chapter 1.

Whenever a situation is perceived or felt to be one which will likely produce an unfavorable gratification-deprivation balance, a group member experiences anxiety and becomes reluctant to express publicly many of his perceptions and feelings. These anxiety feelings may be either consciously or unconsciously experienced. But whether the behavior portrayed is conscious or unconscious in character, many of the perceptions and feelings developed in classroom group situations remain as privately-held experiences by individual members.[18]

At times the class as a whole may be willing to entertain and attempt to deal with individual members' feelings and anxieties; but even when this is true, the group members concerned very often feel that their privately-held perceptions and feelings should not be made public because they will be too threatening to themselves, other members, or the class as a whole.

As a means of allaying anxieties and testing perceptions and feelings being privately held, class members seek to share these perceptions and feelings with other members in whom they feel they can confide. Stated otherwise, classroom group members have a *need* to express and test their privately-held perceptions and feelings with members who will respond actively and supportingly to them. These members are most often those designated as "friends", the intimate associates of a clique or coterie to which they belong.[19]

These relationships which class members seek to develop with one another in order to share and test private perceptions and feelings resulting from participation in classroom situations can be identified as the "friendship" dimension. The question which reveals the status of this dimension at any given time is: Who in the class shares private feelings and ideas with whom?

The Personal Prestige Dimension.—In the process of a classroom group's setting goals for carrying out the learning tasks set by the school curriculum, individual members make various kinds of contributions to the group's progress and welfare. In this way, the

---

[18] Cf. Raymond L. Gorden, "Interaction between Attitude and the Definition of the Situation in the Expression of Opinion." In D. Cartwright and A. Zander, *Op. Cit.*

[19] Provisions for meeting these needs are sometimes institutionalized in a school system in the form of a counselling service.

group discovers the talents and special worth of its different members.[20]

Each member comes to be valued to some degree by the group. Each is accorded personal worth commensurate with his talents, personal characteristics, and contributions to the projects of the group. This acquisition of personal worth represents the group's approval of a member's contributions and the acceptance of him as a person with membership rights, privileges, and responsibilities. For a given period of class operation the status of these relationships can be established by determining who is valued by whom by virtue of contributions made to the welfare of the group and the accomplishment of its goals.

The general approval of him by the group along with the special prestige symbols with which he may be rewarded constitutes the source of one of the most important kinds of psychological rewards a class member receives. "To be wanted", "to belong", "to be approved", or "to win recognition and be honored" is one kind of gratification for which it is common for all people to strive. Without a minimum of this kind of gratification or reward, an individual class member's behavior becomes subject to disorganization, unpredictability, and ineffective problem orientation. Without it a member finds it difficult to continue to take part in the life of the class. In brief, individuals as members of a class have a *need* for being valued by the group.

Briefly stated, the way in which members are related to one another in terms of their being valued by the other group members represents the personal prestige dimension of the classroom group.

The Sex Dimension.—There are two sets of forces that operate so as to make the members of a classroom group relate to one another sexually. First, the society in which members of a classroom group are born and live ascribes or assigns to them different rôles that are based in part upon their sex.[21] People formulate different expectations of one another based on differences in sex and in the classroom these expectations will influence the way in which class members will relate to one another. In almost every kind of situation, formal and informal, in which males and females

---

[20] Cf. C. Heinicke, and Robert Bales, "Developmental Trends in the Structure of Small Groups," *Sociometry*, February, 1953, pp. 7–38.

[21] Kingsley Davis. *Human Society*. New York: The Macmillan Company, 1950, pp. 98–103.

participate, there are socially defined, common expectations as to (1) what activities are appropriate for the different sexes, (2) how females should behave in the presence of males, (3) how females should behave toward other females, (4) how males should behave when with females, (5) how males should perceive and react to other males. These expectations are brought to the classroom and utilized by members, mostly unconsciously, to relate to one another.

Second, members of a classroom group relate to one another in respect to sex because of certain physiological and psychological forces that affect them as individual organisms.[22] That is, each member has physiological and psychological sex needs which move him or her to seek out other class members who have complementary sex needs and/or who might gratify these needs. In everyday life this relationship is often identified as "love".

The relationships class members endeavor to develop between one another (1) to develop group rôles that take account of culturally defined differences in sex and (2) to satisfy psychological and physiological sex drives represent the sex dimension of the classroom group. The status of the relationships along this dimension at any time is determined by establishing (1) which class members are attempting to define participation in the group according to differences in sex, and (2) which members are relating to one another for purposes of gratifying sex needs?

The Privilege Dimension.—From the first of its history, the classroom group includes at least one member, who because of the rôle he or she holds, receives gratifications (or rewards) in the form of deferences, benefits, advantages, and gifts. The classroom group incorporates this one rôle from the very first of its life because the rôle is assigned to it by the larger social organization of which it is a part, *viz.*, the school. This rôle is that of "teacher".

While the person who is competent to fill the position of "teacher" doubtless makes valuable contributions to the work of the group and achieves personal prestige because of these contributions, he also receives some gratifications by virtue of merely filling the rôle. In brief, the teacher from the first of the group's existence is eligible to receive deferences, benefits, and gifts from members.

---

[22] P. Symonds. *The Dynamics of Human Adjustment.* New York: D. Appleton Century Co., 1946, pp. 31–35.

By virtue of holding the position of "teacher", he has certain privileges that no other member has.

As a class has common experiences, builds a culture for itself, and develops a work structure for effective problem-solving, provisions may be made for rewarding members on a privilege basis in addition to those based on their contributions to the achievement of class goals. The group creates and designates rôles that make members holding them eligible for gratifications in the form of advantages, deferences, and gifts. It seems as though individuals, as group members, have a *need* to insure themselves of a constant receipt of gratifications, especially during those times when they are unable or do not have the opportunity to acquire rewards through making contributions to the work of the class. This attempt to control the "flow" of gratifications seems to represent a sort of "hedging" operation by which members attempt to build into the social structure of the class a sort of "guaranteed set of rewards". The more a group builds a "surplus" of needs-gratifying conditions, the more likely class members will seek to work out privilege relationships between one another. The status of the privilege relationships in a classroom group is determined by discovering who has the right to what gratifications by virtue of the rôle they hold rather than by virtue of the contributions they make to the class goals.

Privilege gratifications may be attached to certain rôles in order to insure that they will be filled by persons who are especially competent in making contributions to the work of the group. They may be assigned to a rôle in recognition of the contributions some group member who presently holds it has made in the past to the welfare and needs of the group. Or these gratifications may be assigned to the rôle as a way of "pointing" to the authority figures of the group. Whatever the reasons may be for attaching these gratifications to certain rôles of the group, the point to be recognized and emphasized is that whatever member fills them becomes eligible for receiving the privilege rewards attached to them.

APPLICATION OF THE FRAMEWORK TO CLASSROOM GROUP OPERATION

In the above description of the problem-solving dimension, it was noted that there are two general problems with which classroom groups are continuously confronted. One has to do with

building a kind of class organization that will insure high productivity in the sense of (a) effectively accomplishing the tasks set for the class by the school curriculum, and (b) maximizing the growth or achievement of individual members. The other has to do with maintaining the kind of intra-group conditions that enable members to expend their energies upon class work rather than upon the resolution of inter-member or inter-personal difficulties. These two problems of classroom group operation can be identified as the problems of *group productivity* and *group cohesiveness.*

Class productivity and individual achievement decrease whenever either of these two persisting problems remain unresolved for any length of time. The problem which constantly confronts the teacher is that of diagnosing or analyzing the class to determine (1) which of the two general problems besets the class at any given time and (2) what the character of the relationships are that produce the problem.

Without some kind of conceptual framework to indicate "what kind of things to look for" and "how to look at them", it is not likely that the teacher will be able to bring about a resolution of problems induced by the social structure of the class. When this is the case, class productivity and individual achievement falls and the group more or less "bashes" about with little direction given to its efforts. When confronted with these conditions, a teacher tends to feel extremely threatened and is likely to resort to dictatorial or seductive methods in order to gain more security as the main classroom authority. Under these conditions, class members, including the teacher, seldom understand what is happening to them and it is not surprising that they become highly anxious, reject class work, and some become hostile while others take flight or establish extreme dependency relationships with the teacher.

The systematic use of a framework, such as has been outlined, makes it possible for the educational practitioner to avoid being victimized by the situation just described. For example, the conceptual framework of this article identifies for the educational practitioner the dimensions along which members of a classroom group have a *need* to relate to one another. These dimensions of member relationships constitute the factors relevant to problems induced by the social structure of the classroom group. These different kinds of relationships are "what practitioners must look at" to

discover what the conditions are that create at a given time either one or both of the two general problems of classroom group operation. The conceptual framework also indicates to the practitioner that the productivity of a class, the achievement of its members as individuals, and the member-satisfaction with life in the class is a function of the relationships within and between the seven different dimensions. This is "how" a classroom group must be conceived in order to analyze or diagnose effectively the conditions that upset class productivity and class cohesiveness.

SUMMARY

This article outlines a seven-dimensional framework that can be used to analyze the two general problems of classroom operation with which teachers are constantly confronted. These problems were identified as class productivity and class cohesiveness.

The seven dimensions contained in the framework represent the different kinds of relationships that members of a classroom group have *need* to establish between themselves. They also identify, for the educational practitioner, the factors relevant to operational problems induced by the social structure of the classroom group. In addition, the framework indicates to the practitioner that the productivity of the class, the growth of its members, and general member-satisfaction with life in the class must be conceived as functions of the member relationships within and between the seven identified dimensions.

The conceptual framework enables the practitioner to recognize that whenever the member relationships along and between these dimensions become (1) inadequate for accomplishing the work of the class as set by the school curriculum and (2) unsatisfactory for meeting member's needs, the class will probably experience a drop in productivity and individual achievement. The most effective group arrangement for restoring a high level of productivity is to systematically utilize a conceptual framework (1) to locate the dimensions containing the inadequate or unsatisfactory relationships and (2) for determining how these relationships are affecting class productivity and class cohesiveness.

MENTAL MATURITY VERSUS PERCEPTION ABILITIES IN PRIMARY READING

SISTER MARY JAMES HARRINGTON, S.C.L.

Saint Mary College, Xavier, Kansas

and

DONALD D. DURRELL

Boston University, Boston, Massachusetts

There appears to be great confidence in mental age as a basis for beginning reading. It is often assumed that a mental age of six is the minimum for success in reading, and that later progress is closely allied to mental age.

On the other hand, there is growing evidence that success in reading rests more on specific backgrounds of perceptual abilities, and that these are relatively independent of mental age. Among the background abilities that have been shown to be related to success in reading are the following: visual and auditory perception of word elements (*1*, *2*, *3*, *6*), and ability in phonics (*5*, *6*, *7*, *8*)

It seemed desirable to study the relative influence of mental age and the various perceptual factors on reading, using a single population, and evaluating each factor free from the effects of the others. Such a study should provide light on the following questions: How much of reading success depends upon:

1) Visual discrimination of word elements, required for "look-and-say" reading?

2) Auditory discrimination of word elements, which provides the basis for phonetic reading?

3) Phonics—the attachment of sounds to letters and blends, now coming again into favor.

4) Mental age, the general ability commonly used in educational prognosis?

The second grade at the end of the year seems the most appropriate place to evaluate these factors. Reliable measures were available for the various abilities, the intelligence test used required no reading, and the reading vocabulary was small enough for fairly precise measurement. At this stage, too, school instruction plays a

relatively large part in reading growth as compared to independent reading.

Reading achievement was measured by two tests: an oral reading test and a silent reading word classification test. Both were measures of reading vocabulary rather than comprehension, since the factors studied are presumed to be related to growth in sight vocabulary. The oral reading test was a story which contained 136 unrepeated words from the basal text used by the pupils, *The Cathedral Basic Readers*. The reliability of this measure, obtained by the split-half method and corrected by the Spearman-Brown Formula, was .98. The word classification test was built upon words not necessarily in the basal readers, but in the Dolch list of "First Thousand Words of Children's Reading." It required the child to read the word silently and put it into the proper classification of meaning. This test correlated .80 with the oral reading test.

Visual discrimination was evaluated by showing the children a word on a flash card and having them identify it from memory by drawing a circle around it in a multiple-choice situation. The reliability of this test, computed by the split half-method and corrected by the Spearman-Brown formula was .80. The auditory discrimination test measured the child's ability to notice initial consonant sounds, rhyming at the ends of words, final consonants, and a combination of initial and final consonants in words spoken by the examiner. This test, designed and evaluated by Nason (*4*), required the child to circle words to indicate his knowledge of the stimulus sounds. The split-half, corrected reliability of this test was .90. Ability in phonics was measured by an individual test which asked the child to give sounds of letters, consonant blends, and to show a knowledge of the influence of the final silent "e" on word pronunciation. This test had a reliability of .98, obtained by the split-half method and corrected in the usual manner.

Mental age was measured by the *Otis Quick Scoring Mental Ability Test, Alpha*, Form A. This is largely a measure of ability to understand oral language and to follow directions given by the examiner.

All tests were administered by the first author, and were given to five hundred children in the parochial schools in the Archdiocese of Boston. The children represented a wide range of economic status. All of the schools gave training in auditory and visual discrimina-

TABLE I—RESULTS OF MATCHING 161 PAIRS OF CHILDREN TO DETERMINE EFFECT OF PHONICS ABILITY ON READING ACHIEVEMENT

Variable	Mean of High Group	Mean of Low Group	Difference of Means	S.E.diff	Critical Ratio
Phonics	45.30	17.96	27.34	1.03	26.54
Mental age	97.93	97.46	.47	1.38	.34
Auditory	27.12	26.79	.33	.73	.45
Visual	22.38	21.89	.49	.39	1.24

tion of word elements, and all had a formal or semi-formal program in phonics.

The primary technique for analysis of the influence of each of the factors independent of the others was that of pairing children who were nearly equal on each of three variables, but markedly different in the variable being studied. When the influence of phonics on reading was being studied, for example, 161 pairs of children were found who were equal in mental age, in visual discrimination and auditory discrimination, but one child in each pair was superior to the other in phonics. The technique of pairing always required the children to be within one-half standard deviation on the controlled factors, and more than one standard deviation apart on the independent factor.

Table I illustrates the results of pairing children to determine the effect of phonic ability on reading achievement. This table is typical of the pairings for the various factors. The means of the controlled variables seldom showed a difference of one point of raw score. The difference of the means of the variable being studied was always large, usually more than one and one-half standard deviation.

If the factor being studied has a high influence on reading, the higher children in the pairing should have a higher mean score in reading than the low group. If, however, the factor is of little importance in reading, the two groups should show approximately equal scores in reading.

Table II shows the results of the pairings and the influence of each factor on reading achievement. The factor being studied is indicated by the pairing data being placed in italics.

All three perceptual factors were found to be significantly related

TABLE II—RELATION OF VARIOUS FACTORS TO READING ACHIEVEMENT

Variable	No. of Pairs	Auditory		Visual		Phonics		Mental Age		Reading			
										Scores		Diff.	C.R.
		High	Low	High	Low	High	Low	High	Low	High	Low		
Auditory	106	*34*	*24*	20	20	34	33	93	93	192	174	18	2.7
Visual	134	31	31	*34*	*17*	34	34	95	95	205	173	32	5.8
Phonics	79	28	28	19	19	*38*	*11*	94	94	180	147	33	4.1
Mental age	107	31	31	21	21	31	31	*105*	*87*	192	189	3	.5

Figures in italics indicate the factor being compared to reading.

to reading achievement. Auditory perception made a difference of 18 words in mean reading score; visual perception, a difference of 32 words; phonics, a difference of 33 words. Apparently each has a significant place in the word analysis program, and each makes a distinct and different contribution to progress in developing a sight vocabulary.

Mental age, however, had little influence on success in reading achievement. The high group had a mean mental age of eight years and nine months as compared to seven years three months for the low group; the two groups were equal in the other perceptual abilities; the difference of a year and a half in mental age produced only a difference of three words in reading vocabulary.

A possible explanation for this lack of influence of mental age may be found in the fact that the mental test used is primarily a measure of oral language comprehension. This is the major element measured in most primary grade intelligence tests. It is likely that even the dullest children who come to first grade have a speaking and listening vocabulary adequate for success in first grade reading, and that additional language fluency does not greatly enhance the rate of acquiring a reading vocabulary in primary grades. The minimum estimates of speaking and listening vocabularies of beginning first graders are 2,500 words or above. Since only 500 words are taught in the first grade, and since the more difficult basal readers rarely present more than 1,800 words in the first three grades, it should not be surprising that the relationship between most current primary grade mental tests and reading achievement is low.

A partial correlation study using the entire population of five hundred children confirmed the findings of the paired comparison study, with the minor exception that visual discrimination showed a higher correlation with reading than did phonics. The zero-order correlations with readings for the four factors were as follows: visual discrimination, .64; phonics, .56; auditory discrimination, .54; mental age, .23. Partial correlations were as follows: visual discrimination, .45; phonics, .32; auditory discrimination, .26; mental age, .000.

Since the findings in regard to mental age and phonics are of particular interest, it seemed desirable to repeat the study on a larger population with a wider geographic distribution. Through the coöperation of Sister Mary Brendon, RSM, of Arkansas, Sister Mary Hortense, RSM, of Oklahoma, Sister Catherine Marie, SCL, and Sister Ann Lorraine, SCL, both of Kansas, the study was repeated in the several states.

The same tests and statistical procedures were used with a population of 1,000 children at the end of the second grade. The larger number permitted pairings more than twice the size of the initial Boston study. The findings of this larger study are presented in Table III.

In the data from the midwestern states, the factors relate to reading in the same order as in the Boston study. The one marked difference is in the influence of visual perception, where the difference between the high and low group was only six points, resulting in a somewhat lower influence of the factor in the reading score. The midwestern mental ages ran somewhat higher, permitting a greater difference between the high and low pairs in matching.

TABLE III—RELATION OF VARIOUS FACTORS TO READING ACHIEVEMENT (MIDWESTERN STUDY)

Variable	No. of Pairs	Auditory		Visual		Phonics		Mental Age		Reading			
		High	Low	High	Low	High	Low	High	Low	High	Low	Diff.	C.R.
Auditory	265	*36*	*24*	24	23	40	39	102	101	206	185	21	6.4
Visual	340	32	31	*27*	*21*	41	39	104	104	207	184	23	7.4
Phonics	161	27	27	22	22	*45*	*18*	98	97	189	160	29	6.1
Mental age	283	32	31	24	24	40	40	*122*	*91*	204	198	6	1.6

Figures in italics indicate the factor being compared to reading.

Here a difference of thirty-one months of mental age, or two years and seven months, made a difference of only six words in the mean reading scores.

Referring to the questions stated at the beginning of this article, the findings of these studies are as follows:

1) Auditory and visual discrimination of word elements have high importance in success in acquiring a primary grade reading vocabulary.

2) Phonics instruction is clearly important, having a higher relation to reading achievement than any of the factors studied.

3) Mental age, as measured by the test used, has little influence on success in learning to read.

It is clear from the nature of the pairing that many children who are high in mental ability score no higher than those of low mental ability in the background factors that make for success in reading. This study indicates strongly that specific instruction in phonics and in visual and auditory perception of word elements is essential to success in building reading vocabulary in primary grades.

REFERENCES

1) B. Alice Crossley. *An Evaluation of the Effect of Lantern Slides on Auditory and Visual Discrimination of Word Elements.* Doctor's Thesis, Boston University School of Education, Boston, 1948. (Available on microfilm.)

2) Donald D. Durrell and Helen A. Murphy. "The Auditory Discrimination Factor in Reading Readiness and Reading Disability." *Education* 73: 556–560; May 1953.

3) Helen A. Murphy. *An Evaluation of the Effect of Specific Training in Auditory and Visual Discrimination on Beginning Reading.* Doctor's Thesis, Boston University School of Education, Boston 1943. (Available on microfilm.)

4) Doris E. Nason. *The Influence of Vocabulary Common to Test and Textbook on Primary Reading Test Scores.* Doctor's Thesis, Boston University School of Education, Boston, 1951. Unpublished.

5) Sister M. Dorothy Browne, O. P., *Phonics as a Basis of Improvement in Reading.* Washington, D. C.: The Catholic University of America, 1939.

6) Sister Mary Nila, O.S.F. "Foundations of a Successful Reading Program." *Education* 73: 543–555; May 1953.

7) Joseph Tiffin and Mary McKinnis. "Phonic Ability: Its Measurement and Relation to Reading Ability." *School and Society* 51: 190–192; February 1940.

8) Frank Wilson and others. "Reading Progress in Primary Grades." *Elementary School Journal* 38: 442–449; February 1938.

BOOK REVIEWS

MARGARET NAUMBERG. *Psychoneurotic Art: Its Function in Psychotherapy*. New York: Grune & Stratton, Inc., 1953, pp. 148 $6.75

At the core of Miss Naumberg's latest volume on the use of spontaneous art production in psychotherapy is a three-year study and record of a psychoneurotic patient. A consideration of the psychotherapeutic significance of the spontaneous art productions of this patient takes up by far the largest portion of this volume. Also considered at length are the conclusions drawn from this study. The volume also contains two shorter sections of related material: (1) "The Correlation of the Patient's Rorschach and Other Tests with the Patient's Art Productions" by Adolf G. Woltmann; and (2) "Comments on the Rorschach Test and the Psychotherapeutic Use of Graphic Art" by Zygmunt A. Pietrowski. The book is introduced very favorably with a short preface by Dr. Kenneth E. Appel. The book is an expansion of Miss Naumberg's studies in this field. Of these, the volume on *Schizophrenic Art: Its Meaning in Psychotherapy* has been previously reviewed in this JOURNAL.

The volume includes forty-six illustrations, eight of which are color plates, which show various phases of development as interpreted by the use of these spontaneous drawings. Illustrative of the phases considered are, escape into a pregenital world of fantasy, the family constellation, patterns concerning fear, the growth and development of the patient's ego, the threat of seduction by the father, exploring the meaning of sexuality, marriage and its adjustments, and the patient resolves her fantasy life.

There is also an appendix in two parts: in the first is reviewed the history and development of art in therapy; and in the second is considered the recent literature on clinical and theoretical aspects of spontaneous art in therapy and diagnosis. Some of this material has appeared in previous publications by the author. The material presented, more so than her previous publications, represents a rough measure of a method for evaluating the progress in therapy—ways of detecting, revealing and measuring changes in the patient's personality. Appel in his preface praises Naumberg's therapy as successfully demonstrating the effectiveness of therapy in the area he considers to be refractive; namely, obsessive compulsive condi-

tions. In this particular patient's case, these are compulsive eating, masturbation and marked phantasy life.

As in previous publications, the material here will be useful to therapists using projective techniques in diagnosis as well as therapy. As various publications of Naumberg have appeared, it seems that enough has been written to warrant her writing one more compact volume, tying together the meanings as well as the description of the process of therapy by the use of art productions which is her contribution.

H. MELTZER

Psychological Service Center, St. Louis, Mo.

HENRY CLAY LINDGREN. *Mental Health in Education*. New York: Henry Holt and Co., 1954, pp. 561.

The preface outlines the contents of the book, indicating the plan of organization, and pointing out some things which the author wishes to emphasize. "The chief concern in this book is with the normal child, the normal child who has a few problems (and who has not?), but particularly the normal child who has a healthy need to progress and to develop in ways that are socially and emotionally mature." (p. iv).

The subject matter as shown in the table of contents includes the following in the order given: the mental hygiene attitude in education; development of the child as an individual; motivation; emotional maturity; behavior problems; socioeconomic factors underlying children's behavior; communication; the child and the group; forces operating in children's groups; patterns of acceptance and rejection; three approaches to teaching; the teacher as a person of power and authority; relationship of the teacher to the classroom group; relations with parents; disintegrative influences in education; integrative influences in education; the school as a therapeutic environment; evaluation and diagnosis; the teacher; problems of adjustment. There are five appendixes and an index.

Although 'Mental Health' is stressed in the title, the book is primarily a discussion of problems of education, especially those of teachers. It is essentially a pedagogical text dealing with 'techniques' of teachers, and the 'rôles' of the teachers. Some philosophy of education is brought in to help in the discussions of the various attitudes that teachers may take. These discussions are primarily from the point of view of mental hygiene or mental health. There

are many references at the ends of chapters. They seem to be primarily to pedagogical literature; secondarily to mental hygiene publications.

There are many valuable discussions, especially for the beginning teacher. Many of them seem superficial and too brief, but the wide scope of the book may have forced some of this. The emphasis is on the therapeutic, with the usual lack of recognition upon the very important problem of prevention. Do we not look to the schools to build up the attitudes, habits, interests, etc., that will be the greatest bulwark against the development of abnormal and undesirable character and personality? 'Prevention' does not appear in the index, and does not seem to be essential to the thinking in the text. Perhaps the nearest to it is found in the chapter on Integrative Influences in Education, and in relation to Guidance.

The difficulties of teachers are brought out, but there is much of encouragement for those who are looking forward to teaching as a profession. The status of teachers is better than formerly and indeed quite high in the scale. The salaries are much better. Social recognition has greatly improved. After noting current advantages of teaching, the author says: "Probably all of these factors operate to make teaching a more enjoyable, rewarding, and generally worth-while profession than it was a decade ago." (p. 532).

The index is generally well made (pp. 550–561). The volume follows the consensus of large numbers of educators at the present time. It will probably find a very considerable use as a text. There are a sufficient number of broad generalizations which may be questioned, and thus offer opportunity for extended discussions.

A. S. Edwards

The University of Georgia

Lee J. Cronbach. *Educational Psychology.* New York: Harcourt, Brace and Co., 1954, pp. 628.

During the period in which this book was in preparation, the general expectancy among instructors in the relevant courses was that Cronbach would provide us with a notable text. He has. The author's catholic competency in both psychology and education, his willingness to eschew the widely used formula for writing educational psychology textbooks, and the very evident responsibility with which the book was planned and executed have combined to

produce one of the best of the approximately twenty texts which presently petition the attention of instructors.

The focal worth of this publication can best be suggested by reference to the author's Preface. Here, in a most provocative statement, Cronbach asserts 'four qualities' that an educational psychology must possess "in order to fill the needs of the present day." These are the objectives for the present book and are announced as: "(1) An educational psychology must be relevant to school problems, (2) An educational psychology must be soundly based on research, (3) An educational psychology must be complex enough to do justice to human behavior, (4) An educational psychology must be clear." The self-imposed criteria presented in prefatory pages too often have only a coincidental relationship to the final product; only the contentious will quarrel with the observation that, in this case, the objectives constitute an adequate summation of the text.

The nineteen chapters are organized under five Parts: "Psychology and school problems," "Readiness and its development," "Acquiring ideas, attitudes and skills," "Planning, motivation, and evaluation," and "Emotional learning." Each of these major sections ends in an extensive summary. The end-chapter 'reading lists' are annotated and inviting. Liberal and meaningful use is made of line drawings, figures, and tables. The author's selection of photographs is to be commended; most of the few included appear to extend textual denotations, clarifications, and suggestions.

The technique of including frequent intra-chapter questions, which Cronbach utilized in his *Essentials of Psychological Testing*, is continued here. Inasmuch as one of the reviewers of the earlier text was skeptical of the value of this device, it may be appropriate to note that students evaluating the text at the reviewer's institution are firmly positive on this point.

ARTHUR P. COLADARCI
Stanford University

(*Concluded from Inside Front Cover*)

these charges, as well as information regarding the services of the American Documentation Institute in making available tables, etc., that for economic reasons may be omitted from a printed article, by writing to the Editors or the Publishers.

Double-spacing.—Manuscripts should be typed, written on one side of the paper only, and double-spaced throughout including quotations, footnotes, and bibliographical references.

Footnotes.—Footnotes are to be numbered consecutively beginning with '1', and should be on a separate sheet at the end of manuscript. (Footnotes to tables carry the *, †, and ‡.)

Titles.—Titles of articles should be brief, preferably three to eight words, with an extreme maximum of twelve words.

Type style.—Manuscripts are not to be marked for type style—this is done in the editorial office.

Books and other materials for review, and business communications should be addressed to THE JOURNAL OF EDUCATIONAL PSYCHOLOGY, Warwick & York, Publishers, 10 E. Center St., Baltimore 2, Md.

Subscribers should notify the Publishers of change of address at least four weeks in advance of publication of the issue with which change is to take effect; both the old and new address should be given.

The Publisher desires every subscriber to get all issues to which he is entitled. Each journal is securely enclosed in a sturdy wrapper on which the subscriber's name and address have been printed, and is delivered directly to the Post Office, postage prepaid. Second-class matter is handled less by postal employees than other mail; moreover, if the Post Office is unable to make delivery, a notice to this effect is sent the Publisher and the magazine returned. Consequently, it is doubtful if one journal in many thousands is actually lost in transit.

But after an issue has been delivered to the proper address many things may happen to it—it may be diverted, or misplaced, or borrowed and not returned. For this neither Post Office nor Publisher is responsible. However, a subscriber who does not find a given issue in its assigned place may innocently make a claim of non-receipt. No claim for non-receipt of an issue can be honored unless made within four weeks after arrival of the next succeeding number. In order that a claim may arrive within the time limit it should be addressed to the Publisher—not to an agency.

WARWICK AND YORK *Publishers* BALTIMORE 2, MD.